Implicit Bias Awareness and Intervention Influence on In-service Classroom Teachers Promoting Equity in School Discipline: A Mixed Methods Study

A Dissertation Submitted
to the Graduate School
Valdosta State University

In partial fulfillment of requirements
for the degree of

DOCTOR OF EDUCATION

in Curriculum and Instruction

in the Department of Curriculum, Leadership, and Technology
of the Dewar College of Education and Human Services

March 2021

Nicholas R. Chastain

MAT, Valdosta State University, 2013
BS, Troy University, 2009

This dissertation, "Implicit Bias Awareness and Intervention Influence on In-service Classroom Teachers Promoting Equity in School Discipline: A Mixed Methods Study," by Nicholas Robert Chastain, is approved by:

Dissertation Committee Chair	_____ James L. Pate, Ph.D. Professor of Curriculum, Leadership, and Technology
Dissertation Research Member	_____ Lantry L. Brockmeier, Ph.D. Professor of Curriculum, Leadership, and Technology
Committee Member	_____ Don Leech, Ed.D. Professor of Curriculum, Leadership, and Technology
Associate Provost for Graduate Studies and Research	_____ Becky K. da Cruz, Ph.D., J.D. Professor of Criminal Justice
Defense Date	_____

FAIR USE

DUPLICATION

I authorize the Head of Interlibrary Loan or the Head of Archives at the Odum Library at Valdosta State University to arrange for duplication of this dissertation for educational or scholarly purposes when so requested by a library user. The duplication shall be at the user's expense.

Signature _____

I refuse permission for this dissertation to be duplicated in whole or in part.

Signature _____

ABSTRACT

Discipline disproportionality is a topic studied and discussed extensively in the United States. Schools and districts across the nation have long sought pragmatic solutions to this long festering problem. Inequity in school discipline is considered by many as the main cause of the achievement gap and a host of other negative student outcomes. Scholars have studied the construct of implicit bias for decades, yet many still consider the topic controversial. The purpose of this study was to examine how teacher perception of the topic (equity in school discipline) changed over time by participating in the study.

This study used the sequential explanatory mixed method design. A total of 60 in-service classroom teachers completed the pretest and posttest Teacher Multicultural Attitudes Survey (TMAS). Teacher scores on the (TMAS) and the Race Implicit Association Test (IAT) were examined and analyzed. The qualitative portion of the study examined the responses of nine teachers who completed the quantitative portion of the study.

Results from the quantitative portion of the study were statistically nonsignificant. Several issues of practical significance were identified. Teacher repeated measure scores on the IAT indicated a slight change in preference from White to Black skin. In contrast, teacher pretest and posttest scores on the TMAS indicated teachers had less awareness of and sensitivity to multicultural issues in their classroom. Teacher interview data, however, indicated teachers seeking to improve their efficacy are willing to discuss and address this problem.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my dissertation committee. Dr. Leon Pate, dissertation chair, Dr. Lantry Brockmeier, my researcher, and Dr. Don Leech, my reader. Dr. Pate, when you agreed to serve as my chair, you gave me a sense of hope during a seemingly hopeless time. Your encouragement and support while I was at Duke with my wife is something I will never forget–thank you. Dr. Brockmeier, you have the patience of Job. Your unrelenting attention to detail is an asset to every student you teach and support. I would be woefully unprepared to make decisions based on data and research without your persistence. Thank you, Dr. Leech, for agreeing to serve as my reader sight unseen. Your pragmatic curiosity during my proposal defense saved me from a lot of trouble later in the process.

I would like to thank Dr. Frederick Smith, a former principal, mentor, and colleague. You gave me the confidence and freedom to seek pragmatic solutions to sensitive issues. I would like to thank Brian Knighton, friend and colleague, for allowing me to grieve over racial inequity inherent in our society. I doubt I would have had the courage to face my own shortcomings if you had been unwilling to help. I would also like to thank Dr. Richard Schmertzing. Your course on qualitative research literally changed my life. I sincerely doubt I would have taken on this topic without you. I know I would not have had the passion or perseverance necessary to finish this dissertation had I not met you. I would also like to thank Dr. Marni Kirkland for believing I could grapple with this topic. Your consistent push to end inequity in education is inspiring and hopeful.

I would like to thank my beautiful wife Bonnie for her belief I could see this through. You pushed and prodded me in ways no one else could. Your mixture of tough love and patient persistence is something I will always be thankful for. The long nights at Duke, the weekends you had to spend alone, the financial sacrifices you made, the list goes on. I owe you a great deal for reaching this point in my career and I will be eternally grateful for your sacrifice and support. I would like to thank my brother, Chris Chastain, for helping me to relax (sometimes). Your ability to smile in the face of adversity and put things in perspective is what makes you one of the best leaders I have met. You have a gift for inspiring others to be their best, and I am proud to call you brother. To my mom and Dad, Louie and Cathy Chastain, thank you for letting me use the beach house those few times I couldn't find the words to write. Thank you for encouraging me to read when I was young, and having patience when I was lost. I love you. Finally, I would like to thank all of the friends, family, and colleagues who supported me through this journey.

DEDICATION

This dissertation is dedicated to Bonnie. Thank you for loving me, believing in me, keeping me honest, and making me think. I love you.

This dissertation is also dedicated to all of the students waiting to meet their favorite teacher–and the teachers working tirelessly to be *that* person.

Chapter I

INTRODUCTION

Discipline disproportionality is a problem plaguing many schools in the nation– especially the Southeastern United States (U.S. Department of Education, 2016a). The U.S. Department of Justice and Education defines discipline disproportionality as the discipline of a racial or ethnic subgroup of the students in a school or district at a higher rate than the total student population of the school or district (2014). The U.S. Department of Education's Office for Civil Rights releases data annually that describe and measure equity in school discipline in the United States (2016b). Losen, Hewitt, and Toldson (2014) reported the number of schools able to reduce or eliminate discipline disproportionality are few. The Civil Rights Data Collection (U.S. Department of Education, 2016c) reported African American students were suspended at nearly four times the rate of European American students in the United States during the 2013-14 school year. The U.S. Department of Education (2016a) claimed students who were removed from the learning environment for discipline reasons often face negative consequences possibly affecting them long after they become adults. Students who were suspended were more likely to drop out of school and never graduate, fell further behind academically, and were more likely to be suspended again. Losen et al. (2014) reported there is no systemic policy or guideline proven to reduce discipline disproportionality. However, schools should be able to effectively eliminate this discriminate practice. Gregory, Skiba, and Noguera (2010) assert there is no current or historical research

definitively helping schools reduce discipline disproportionality. Gregory et al. alleged the problem is getting worse.

Implicit racial bias is an emerging research topic in the field of educational discipline disproportionality. In 2010, van den Bergh, Denessen, Hornstra, Voeten, and Holland conducted a landmark study investigating if implicit racial bias might play a role in the ethnic achievement gap of students in the Netherlands. Although the study was relatively small and conducted in a geographically isolated region, the findings were strong enough to provoke future research in other settings and locations. Van den Bergh et al. found when teachers self-reported their racial biases most indicated they did not have any. These same teachers then took the Implicit Association Test (IAT) on ethnicity. The results of this assessment indicated most teachers had implicit racial bias against ethnic minority students. Multi-level analyses were conducted comparing self-report measures, student achievement data, and implicit bias score. Van den Bergh et al. reported implicit racial bias, rather than self-reported racial bias, was a far more accurate predictor for teacher expectations of ethnic minority students. The authors also suggested teachers with implicit bias think ethnic minority students were less intelligent, and the achievement gap was wider in classrooms of teachers who displayed strong implicit bias.

Conceptual Framework

Gregory et al. (2010) suggested classroom level characteristics and differences are largely responsible for discipline disproportionality. Anderson's (2015) optimal resource theory "adopts a pragmatic approach that focuses on incremental rather than systemic change by examining micro-policies and practices at the local education level" (p. 27). Anderson further explains it is difficult to judge if systemic change is effective

due to inherent differences in implementation efficacy. When attempting to improve student outcomes (better grades, increased attendance, etc.), school and classroom level decisions and practice can be powerful predictors (Anderson, 2015).

Tenenbaum and Ruck (2007) found most teachers favored European American students in the United States. They came to this conclusion after conducting four meta-analyses examining teacher expectations, office discipline referrals, and both positive and negative speech along racial differences. McKown and Weinstein (2008) in a similar study, found teacher expectations were significantly lower for African American students than for European American students. The study also found in classes where teachers had the lowest expectations for African American students the achievement gap was highest. McKown and Weinstein termed these classrooms low bias and high bias respectively. The van den Bergh et al. (2010) study found teacher implicit bias seemed to play a strong role in the size of the ethnic achievement gap in the Netherlands. The study indicated teacher's self-report or explicit bias had no bearing on achievement gap size. The possibilities for future research these three studies created concerning racial inequity in education are still being discovered today. Implicit bias is a term becoming well known and a concept studied more often as researchers try to understand how it may affect teacher decisions in the classroom, the achievement gap, and discipline disproportionality (McIntosh, Ellwood, McCall, & Girvan, 2017).

Greenwald, Banaji, and Nosek (2015) developed the Implicit Association Test (IAT) in 1995. They were interested in measuring the difference between self-reported attitude measures and more subtle and difficult to measure implicit attitudes. Implicit attitudes or biases were thought to be more difficult to measure because most individuals

are not aware of their implicit biases (Greenwald, Nosek, & Banaji, 2003). The IAT measures implicit bias by using timed computer administered categorization tests based on strength of association (Greenwald, Poehlman, Uhlman, & Banaji, 2009). From a psychological perspective, the difference in explicit and implicit attitudes is based on time and conditions (Gawronski & Bodenhausen, 2006). Explicit attitudes are those evident when a person has the time and cognitive resources to make an informed decision. While implicit attitudes are reported to be the automatic decisions used in snap judgements when time is limited or stress is high (Gawronski & Bodenhausen, 2006).

Smolkowski, Girvan, Mcintosh, Nese, and Horner (2016) report this distinction is important in the field of education because teachers inherently make snap decisions multiple times a day. Smolkowski et al. theorized most teachers in the United States are not overtly racist. They believe most teachers go out of their way to appear not to be racist (due to societal expectations). They further believe most teachers do have implicit racial biases against African American students. This combination of attitudes was termed the aversive racism theory (Dovidio & Gaertner, 2000). Smolkowski et al. believe aversive racism leaves teachers at risk for administering discipline in racially disparate ways.

Statement of the Problem

Discipline disproportionality has been an issue in the American educational system for decades (U.S. Department of Education, 2016a). The Children's Defense Fund (1975) reported, in 1973, African American students were suspended at a rate over three times more than their European American peers. The Children's Defense Fund's classic work *School Suspensions: Are They Helping Children* (1975) suggested this trend

began with the United States Supreme Court's landmark decision desegregating all public schools (Brown v. Board of Education, 1954). Indiana University's Equity Project (2016) defines disproportionality "as the over or under-representation of a group in a category that exceeds our expectations for that group, or differs substantially from the representation of others in that category" ("What is Disproportionality", para. 1). Even with a clear definition of discipline disproportionality and a strong mandate from the Federal Government, most schools fail in their efforts at reducing discipline disproportionality (U.S. Department of Education, 2016b). Discipline disproportionality is a complex problem with no easy answer (Gregory et al., 2010).

French and Ernsthausen (2014) speculated improved school climate and less punitive discipline practices could help reduce the problem of discipline disproportionality in Georgia. Freeman and Steidl (2016) indicated schools and districts need to do more in an effort to end the travesty of disproportionately suspending African American males. Losen et al. (2014) recommended examining school and classroom level differences as guides that may help reduce discipline disproportionality.

Teacher decisions concerning discipline can vary a great deal from classroom to classroom (Gregory et al., 2010). This disparity may be a contributing factor to what Gregory et al. (2010) coined *differential behavior*. According to Skiba and Peterson (2003), the underlying reason for writing a discipline referral often varies according to race. African American students tended to be written up for subjective reasons (student incivility and loitering) while European American students were more often written up for observable violations (vandalism and smoking). Skiba and Peterson further proposed this variation in discipline practice indicated discipline disproportionality starts at the

classroom level. Cultural differences between a predominantly European American teaching staff and African American students, subconscious biases, and low expectations for African American student achievement are all systemic explanations for discipline disproportionality (Skiba & Peterson, 2003).

Gregory et al. (2010) indicated teachers have a great deal of flexibility in determining what behaviors constitute an office referral and what consequences are doled out. The fact African American males are often issued office discipline referrals (ODR) for subjective reasons (Skiba & Peterson, 2003) gives credence to Krezmien, Leone, and Achilles' (2006) assertion that classroom level differences of discipline disparity occur along racial lines. They indicated discipline disparity for African American students was increasing despite policies and mandates created to reduce or end this practice.

Purpose of the Study

Gregory et al. (2010) suggested discipline disproportionality is helping proliferate the achievement gap between African American and European American students. Nance (2017) believes discipline disparity for African American students is the largest contributing factor to our current and growing school-to-prison-pipeline. The problem has continued to grow for decades and has become both entrenched and systemic. Simmons (2015) suggested it has risen to a level indicating a violation of civil rights. Gordon, Piana, and Keleher (2000) found discipline disproportionality can also lead to exclusionary factors negatively effecting academic achievement and helping increase the probability African American students will drop out of school. Losen et al. (2014) agreed exclusionary factors like suspension or expulsion, when viewed through the lens of

disproportionality, negatively affect student learning outcomes and contribute to the achievement gap.

Georgia educators and institutions have struggled with finding pragmatic ways to reduce discipline disproportionality for decades. The most recent data available indicate Georgia's discipline ratio is increasing (U.S. Department of Education, 2016a). A new approach is needed to help schools and teachers understand why this is occurring. Teachers need interventions to become better educators for all of their students. In the past decade, research indicates there are strategies effective in other settings and locations (Devine, Forscher, Austin, & Cox, 2012; Forscher, Mitamura, Dix, Cox, & Devine, 2017; Pepis, 2017). This study was designed to better understand if making in-service teachers concerned about equity in school discipline aware of implicit bias and giving them strategies through an intervention designed to overcome bias are effective at reducing their bias.

Devine, et al. (2012) developed an intervention purposefully designed to help individuals overcome implicit racial bias. Devine et al. believed implicit racial attitudes and their unintended consequences were much like thoughtless bad habits (biting fingernails etc.) and could be altered by motivated individuals. The intervention educates individuals about the topic of implicit bias through short vignettes that describe what implicit racial bias is and how it negatively effects African Americans in our society. The intervention then teaches participants various strategies to practice that will lessen implicit racial attitudes and (hopefully) their often-unintended negative consequences. In her study, participants were randomly assigned to either the control or intervention group. The control group did not have access to her intervention during the study. Devine et al.

found the intervention effective at increasing participant concern about implicit racial bias and lowered implicit racial bias as measured by the IAT over time using a General Linear Model (GLM). Participants in the intervention group (n = 53) lowered their IAT scores significantly more than control group (n = 38) participants in repeated IAT scores, β = -.19, $t(88)$ = -2.82, p = .006, $\Delta R2$ = .081.

The literature and theories on why schools continue to remain out of compliance with federal guidelines concerning discipline disproportionality is substantial (Losen et al., 2014). Smolkowski et al. (2016) believe implicit racial bias is a pervasive, systemic cause fueling this decades old problem. The purpose of this study was to investigate whether an intervention designed to make teachers aware of and reduce their implicit bias is successful. Issues surrounding race are inherently sensitive subjects. Howell, Gaither, and Ratliff (2014) found participants became defensive when presented with Race IAT implicit bias results indicating white preference. The Race IAT uses the colors of white/black rather than the racial monikers African American or European American. Further, participants who rated themselves as pro-black in their explicit measure assessment and received IAT results indicating a strong preference for white were the most defensive. These findings indicate that awareness of their implicit bias is not enough. Plant and Devine (2009) found participants who reported they wanted to achieve results indicating racial neutrality were able to improve their implicit bias scores. For this to occur, participants had to be made aware of their implicit bias and taught strategies designed to help them reduce their racial bias. Devine et al.'s (2012) interactive intervention was designed to reduce prejudiced habits. The purpose of this study was to

specifically target in-service teachers interested in reducing their implicit bias and find if they are able to successfully achieve that goal with the intervention.

This study was grounded in the belief that teacher decisions at the classroom level are ultimately driving discipline disproportionality. Finding ways to help teachers understand their role in this process without making them defensive was the goal. Using in-service teacher volunteers concerned about equity in school discipline for the study was intentional and aimed to reduce implicit bias attitudes. The scope of this study is very limited due to the small population size and sample. The quantitative portion of this study was conducted completely online in May 2020 due to shelter in place orders and school closures caused by the COVID-19 pandemic. The qualitative portion of this study (also online) was conducted in August 2020. By utilizing only volunteer in-service teachers, who were unable to teach their students or tell them goodbye at the end of the year, participant bias is a threat to the overall generalizability of the results of this study.

Research Questions

The following research questions guided this study:

1. Is there a significant difference between pretest and posttest scores on Ponterotto's (1995a) Teacher Multicultural Attitude Survey (TMAS) by the control group and experimental group?
2. Is there a significant difference among participants who score low or high on the pretest TMAS on the final IAT score?
3. Is there a significant difference in participant repeated measure scores on the Implicit Association Test (IAT)?

4. In what ways do the interview data of licensed teachers about their views on the importance of awareness of implicit bias in classrooms provide an explanation for any quantitative results from the IAT or TMAS?

Significance of the Study

Implicit bias research is an emerging topic in the field of education–specifically in the area of discipline disproportionality. Bottiani, Bradshaw, and Mendelson (2017) recently called for research on identifying interventions designed to help teachers and leaders understand the role implicit bias may play when interacting with African American students. They found African American students by in large, had a negative perception of school equity and did not feel as welcome in their school as European American students. The negative affect was even larger in schools with a high ratio of disproportionately suspending African American students. Previous studies by Devine, et al. (2012) and Pepis (2017) found it was possible to reduce implicit bias in undergraduate psychology students and pre-service teachers respectively. This study added to the body of literature by exploring the effectiveness of this intervention at reducing implicit bias in licensed in-service teachers.

Methodology

This study employed the mixed methods explanatory sequential design. The intervention was closely aligned to the previous studies conducted by Devine et al. (2012), Forscher et al. (2017), and Pepis (2017). The quantitative portion of this study sought to determine if implicit bias could be reduced through the use of an intervention. Participants ($N = 78$) were randomly assigned to either the control group ($n = 39$) or experimental group ($n = 39$) utilizing Qualtrics software. The control group did not take

the IAT, the experimental group did. Forscher et al. theorized participants would be less defensive if they were educated about the IAT and understood what it is measuring before using the instrument. A one-way analysis of covariance (ANCOVA), a one-way analysis of variance (ANOVA), and a one-way repeated-measures ANOVA were used to determine any significant differences between the groups or over time. The qualitative portion of the study consisted of participant interviews. Participants from both the control and experimental group were included. The results from this portion of the study gave context and clarity on perceived intervention effectiveness and thoughts on whether this process might help reduce discipline disproportionality in their school or classroom.

Organization of this Study

This dissertation is comprised of five chapters. Chapter 1 presented a conceptual framework, problem statement, purpose of the study, significance of the study, and brief description of the proposed methodology for this study. Chapter 2 synthesizes the literature on discipline disproportionality and its effect on African American students, implicit bias research, the effects it can have on interracial interactions, a review of the intervention, and the instruments employed in this study. Chapter 3 provides a detailed view of the research questions, the quantitative and qualitative design plan, population utilized in the study, instrument descriptions, data collection procedures, and analysis. Chapter 4 included the findings of this study. Chapter 5 is comprised of a discussion of the findings and possible ideas on further research.

Chapter II

LITERATURE REVIEW

This chapter will review the five related areas of literature and research that comprised this study. The first section contains the definition of discipline disproportionality, including both the methods and rationale governmental agencies and educational researchers use when calculating disproportionality. It was important to understand these differences because various methods of calculation are used to measure disproportionality for a variety of reasons (risk ratio, absolute rate by subgroup, odds ratio). The second section explored how disproportionality negatively effects African American students. The third section covers previous efforts to address equity in school discipline. The fourth section introduces and discusses the psychology behind implicit bias. The fifth section reviews the literature on the Implicit Association Test (IAT), Devine et al.'s 2012 intervention, and previous research.

The purpose of this study was to investigate whether in-service teachers who are interested in learning about and reducing their implicit bias are able to do so utilizing a revised prejudice-habit breaking intervention by Devine (Forscher, 2016). Recent studies indicated some success with this type of intervention and have called for further research as a means to help reduce or eliminate discipline disproportionality in American schools (Glock & Klapproth, 2017; McIntosh, Girvan, Horner, & Smolkowski, 2014; Peterson, Rubie-Davies, Osborne, & Sibley 2016). This study aimed to contribute to this particular type of research in a setting with participants who have historically struggled in the area

of disciplining African American students disproportionately (French & Ernsthausen, 2014).

Defining Discipline Disproportionality

The Children's Defense Fund (1975) first reported on the phenomenon of discipline disproportionality over 4 decades ago in their seminal work, *School Suspensions: Are They Helping Children*. Discipline disproportionality is loosely defined as the over or under-representation of a racial or ethnic subgroup when calculating the discipline ratios for schools, districts, or states (U.S. Departments of Justice and Education, 2014). Skiba, Michael, Nardo, and Peterson (2000) investigated discipline disproportionality in three areas: gender, race, and socioeconomic status. They found African American males were disproportionately suspended more than their European American peers and were more likely to be issued an office discipline referral (ODR) for subjective reasons. Interestingly, they found gender, F (3, 10,776) = 310.56, p <.001, and race, F (3, 10,776) = 165.35, p < .001, were consistent predictors of disproportionality while socioeconomic status had a minimal overall effect. Skiba et al. pointed out administrators did not punish African American students differently than other subgroups committing similar offenses. They posited disparate discipline practice was instead caused by systemic bias at the classroom level. Skiba et al. made clear they were not implying or stating overt racism was to blame for discipline disproportionality, but called for more research in this area 20 years ago. Remarkably, McIntosh et al. (2017) specifically warn against using the term "racism" when discussing equity in school discipline with teachers.

Discipline disproportionality can be measured a number of ways. McIntosh et al. (2017) believe schools and districts should calculate school discipline with two different formulas: a risk ratio and absolute rates by subgroup. They recommend using two types because using just one may misrepresent possible problems or accomplishments. The risk ratio is calculated by taking the percentage of students receiving an ODR and dividing that number by the percentage of another group receiving an ODR (males, females, students with disabilities, etc.). The absolute rate is calculated by dividing the number of students who received an ODR in a subgroup by the total number of students in that subgroup. Once this is accomplished for all subgroups in a school or district the local education agency (LEA) can compare their ratio to national, state, and regional norms. Using these two metrics allow schools and districts to plan for and understand if they are meeting federal regulations designed to reduce discipline disproportionality.

Brown and Steele (2015) utilized The Relative Rate Index in their study to quantify school discipline disproportionality. This method is endorsed by the U.S. Department of Justice. The results from a Relative Rate Index analysis indicates a ratio of disproportionality. This is accomplished by taking the number of discipline incidents for African American students and dividing it by the total number of African American students at the school, then multiplying that number by 100 (African American student suspension rate). Follow the same formula for European American students (European American student suspension rate), then divide the African American student suspension rate by the European American student suspension rate and the rate of discipline disproportionality is produced. A result of 1.0 would indicate racial discipline equality. A result of 2.0 would indicate African American students were suspended at twice the

rate as their European American peers. A result of 3.0 would indicate African American students were suspended at three times the rate as their European American peers etc.

Krezmien et al.'s (2006) study provides an additional formula useful in understanding the ways discipline disproportionality can be calculated. Their study examined multiple variables using advanced multivariate statistical procedures to determine disproportionate discipline rates and possible predictors of discipline incidents. To calculate the rate of suspensions, Krezmien et al. took the total number of suspensions per 1,000 students and the number of students suspended per 1,000 students to create two separate data sets. Krezmien et al. then used race as a predictor utilizing logistic regression techniques. Finally, they utilized race and disability (over 40 categories) to analyze discipline disproportionality. The rate of suspension calculation adds context to the previously mentioned Relative Rate Index analysis. Utilizing this technique would calculate a school or district's overall *culture* of discipline. Krezmien et al. found African American students were suspended at a rate significantly higher than their European American peers. They also found the rate increased from year to year. Okonofua and Eberhardt (2015) reported most disproportionality calculation methods do not take into account the number of times a student receives an ODR during the school year. Krezmien et al. attempted to include that data in their study. Unfortunately, disaggregated data from Maryland only had one option, suspended or not suspended. Okonofua and Eberhardt claim in national data sets, if multiple suspensions were included in disproportionate ratios, the rates would be even worse for African American students.

Freeman and Steidl's (2016) analysis of discipline disproportionality in Georgia schools accounted for multiple suspensions by creating a suspension imbalance ratio. This was accomplished by comparing the proportion of African Americans at a school or district to the proportion of all suspensions–even multiple per student–for the same subgroup. Other variables utilized included measures on school segregation (dissimilarity, entropy, exposure, and isolation indexes) and control variables for school demographics, organization, and resources. They predictably found African American students were disproportionately disciplined in most schools and districts. The control variables, however, yielded some significant findings. Freeman and Steidl reported segregated schools were less likely to have an imbalance ratio for African American students. The exposure index variable, alternatively, suggested when the percentage of African Americans attending a school was higher than the racial composition of the school district, the suspension imbalance ratio increased for African Americans. Furthermore, when the isolation index variable for African American student percentage at a school was *lower* than district composition the suspension ratio decreased for African Americans. Freeman and Steidl claim these findings indicate as schools integrate, the suspension imbalance for African Americans increases. They allege this pattern of exclusionary discipline on the African American population in both integrated and exposure schools may represent a pattern of re-segregation. While this allegation is clearly hypothetical and controversial, the suggestion of subtle systemic inequities in how discipline decisions are made, and the negative effects these decisions have on the African American population is disturbing.

Effects of Discipline Disproportionality on the African American Subgroup

Noltemeyer, Ward, and Mcloughlin's 2015 meta-analysis illustrated how exclusionary discipline negatively effects both dropout and achievement rates for all students ($Q(24) = 16{,}079.26$, $p < .001$). Their findings indicated out of school suspension (OSS) had more of an effect than in school suspensions (ISS) on academic achievement and dropout rate. Although their data was not disaggregated to allow for racially disparate outcomes, the results are still troubling. It stands to reason if exclusionary discipline practices were determined to negatively affect student outcomes, and African American students are disproportionately exposed to exclusionary discipline, then the fragile African American subgroup will suffer further deficits in achievement (Gregory et al. 2010). Gordon et al. (2000) wrote a scathing report nearly 20 years ago concerning disproportionate discipline. They reported exclusionary discipline practice and *zero tolerance* policies would increase African Americans involvement in the school to prison pipeline (STPP).

Nance's (2016) report on the STPP analyzed restricted data from the National Center for Education Statistics (NCES). He found disproportionate exclusionary discipline practice did seem to put African American students at a higher risk of being referred to law enforcement or being arrested. Although African Americans only comprised 16% of the student population, they accounted for 27% of law enforcement referrals and 31% of arrests (Nance, 2016). Nance (2017) later reported school surveillance tactics and zero tolerance policies were found to be more intense and obtrusive in schools with majority African American populations. He called for

legislation designed to reduce this problem. Nance (2017) believes intense surveillance and zero tolerance policies,

> especially when applied disproportionately to students of color, harms students' interest, delegitimizes the educational process, perpetuates racial inequalities, weakens trust in government institutions and processes, skews minorities' perceptions of their standing in our society, and sends harmful messages to members of all races that students attending majority white schools enjoy great privileges and have superior privacy rights. (p. 831)

Nicholson-Crotty, Birchmeier, and Valentine (2009) found disproportionate discipline practice led to higher arrest rates and juvenile justice system involvement for African American youth in Missouri. They alleged African American students were more likely to face punitive punishment than European American students who commit similar offenses. They asserted disproportionality still occurs after accounting for environmental factors such as poverty, urban density, and parental employment status. The literature on the damaging effects of discipline disproportionality (STPP, juvenile justice involvement, arrest record) on the African American population is substantial (Bleyaert, 2009; Fisher, 2011; Gass & Laughter, 2015; Irby, 2013; Monahan, Vanderhei, Bechtold, & Cauffman, 2014). Bell's (2015) qualitative study found African American males relished going to school. The participants in his study enjoyed socializing with friends, exercising, and learning new things. When Bell started collecting data on student perception of teachers, however, the enthusiasm and joy all but disappeared. African American males stated teachers seemed to hold them to a different standard (behaviorally) than their European American peers. Interestingly, many participants

described teachers as nice and helpful–until they broke a rule. Once a rule was broken, teachers were overwhelmingly characterized as ruthless disciplinarians. Disciplined students believed teachers were willing to overlook or marginalize disruptive behavior by European American students while African American students were given an ODR and sent out of the classroom with no questions asked. Participants suggested teachers preferred using exclusionary options when disciplining African American students.

Beck and Muschkin (2012) hypothesized exclusionary discipline was closely correlated with the racial achievement gap in North Carolina. They analyzed the administrative data for all seventh-grade students in the state. This data included disaggregated records for discipline and achievement. Beck and Muschkin included a compelling graph comparing reading mastery and disciplinary infractions by race. The European American subgroup scored over 80% in reading mastery. The African American subgroup lagged far behind at 53%. In contrast, over 30% of African American students were issued an ODR–compared with only 14% of European Americans. Beck and Muschkin found discipline and achievement outcomes were strongly correlated. Their analyses suggested disciplinary events had a compounding effect on academic achievement. The more a student was disciplined, the further behind academically they became. Students struggling academically were given more ODR's. Additionally, African American students were disproportionately retained in grade for poor academic achievement causing those students to become old for grade. The data in this study indicated being old for grade was also a strong predictor for increased disciplinary events and lower academic achievement. In a similar study, Behnken (2014) found African American students identified with ADHD had higher teacher ratings for

misbehavior, lower test scores, more exclusionary discipline events, and were more likely to enter the STPP. In fact, Behnken describes a more sinister moniker for this phenomenon: The *Cradle* to Prison Pipeline. Behnken argued teacher perception of African American children at an early age can cause long lasting consequences for the student. Low academic achievement and exclusionary discipline are both significant predictors of juvenile justice contact and eventual incarceration as an adult, explained Behnken. Shollenberger's (2015) analysis of the National Longitudinal Survey of Youth support Behnken's claims. Shollenberger found African American males had higher rates of suspension and incarceration and lower achievement scores. Interestingly, when Shollenberger controlled for suspension, these disparities in achievement and arrest diminished significantly for African American students. Shollenberger also found *any* student suspended 10 days or more was significantly more likely to be arrested or drop out of school. Unfortunately, African American males made up a disproportionately large percentage of this at-risk group (Shollenberger, 2015).

Sullivan and Bal's (2013) research on special education disproportionality yielded some interesting results applicable to this study. African American students were routinely over identified for special education and under identified for gifted services across the nation. They argue discipline disproportionality may be partly to blame. They believe exclusionary discipline causes African American students to miss valuable seat time and instruction, thereby falling further behind their European American peers academically. African American students were identified as special education eligible nearly three times more than European Americans. Sullivan and Bal contended students labeled as special education eligible often face decreased teacher expectations, segregated

learning environments and negative teacher perception. Vanderhaar, Munoz, and Petrosko (2014) report African American students (with or without disabilities) were three times more likely (13.1%-3.8%) to be placed in an alternative school for disciplinary reasons at least once in their educational career. As is common when exploring discipline data, subsequent placements are not reported. Placement in the alternative school, considered by many a segregated learning environment, was a strong predictor of subsequent arrest and lower test scores (Vanderhaar et al., 2014).

Fiester and Gibson's (2015) study supported Sullivan and Bal's (2013) hypothesis. Fiester and Gibson recognized race was not in and of itself a predictor for academic achievement (end of the year oral reading fluency [EOYORF] scores). They did find, however, ODR's and subsequent exclusionary discipline practices along with beginning of the year oral reading fluency (BOYORF) scores were a significant predictor of poor academic growth and achievement ($F(4, 512) = 210.95$, $p < .001$). They contend schools need to limit exclusionary discipline practice–especially for minority students already struggling academically. The compounding effect of missing class from year to year does nothing but exacerbate both the achievement and discipline gap for African American students (Arcia, 2006; Balfanz, Byrnes, & Fox, 2014). In a similar study, Hughes and Kwok (2007) reported African American students were more likely to have a negative relationship with their teacher, and teachers were more likely to have a negative relationship with the parents of African American students. These negative relationships were strongly correlated with poor gains in reading as measured by the Woodcock-Johnson III Test of Achievement (2007). While discipline disparity was not a variable used in the statistical analysis, both peers and teachers indicated African American

students were more aggressive and disruptive than European American or Hispanic students. Teachers also seemed to communicate less often with African American parents and the message was often negative. Hughes and Kwok (2007) contend these negative relationship constructs may cause a disproportionate number of African American students and parents to disengage from the educational process. Mortenson's (2018) analysis found many teachers in the National Center for Education Statistics 2002 Educational Longitudinal Study had significantly lower expectations and perception of ability and work ethic for African American students than European American students. Mortenson proposed the data indicated many teachers in the study had implicit bias against African American students. Mortenson then compared teacher perception of students to standardized math assessment scores and found a strong correlation with the racial achievement gap. African American students, all else being equal, scored nearly 5 points lower on the standardized math assessment used in the study (Mortenson, 2018). Losen et al. (2014) cited many of these possible negative outcomes in their research brief on discipline disproportionality. They argued African American youth are harmed in numerous ways by exclusionary discipline. This process is costly for both the African American community and the nation. Poor learning outcomes, STPP, and reduced potential lifetime earning power were all more likely scenarios for African American (than European American) youth. Schools and districts have tried to ameliorate this devastating cycle in a number of different ways–only a rare few have been successful (Losen et al., 2014).

Current Attempts to Reduce Discipline Disproportionality

Reducing discipline disproportionality in our nation's public schools is a top priority for most districts (Losen et al., 2014). One of the more popular trends is cultural responsiveness professional development. Blitz, Anderson, and Saastamoinen (2016) recently conducted a mixed methods study at an elementary school in the Northeastern United States. The school's demographics had changed drastically over the last 30 years. The researchers reported suspension rates were up, test scores were down, and African American students were disproportionality disciplined. The school district hired a successful consulting group to deliver a cultural responsiveness professional learning workshop. The district informed Blitz et al. teachers were required to attend the training. The results from this study yielded some surprising insights. Blitz et al. found, by in large, teachers at this school cared deeply for their African American students. Many teachers reported a level of sadness and helplessness as they described the home lives some of their African American students endured. In stark contrast, they rated their African American students as more disruptive, aggressive, and apathetic than European American students. When questioned about the training, however, nearly all of the responding teachers indicated they were offended by the insinuation of being culturally insensitive. Blitz et al. emphasized many of the teachers reported they were *colorblind* and treated all of their students the same, regardless of race and socioeconomic status. Sue et al. (2007) categorize color blindness as a micro invalidation–an unconsciously deployed microagression thought to marginalize the psychological effects implicit bias may have on African Americans.

Hartmann, Croll, Larson, Gerteis, and Manning (2017) believe individuals who identify as colorblind are not necessarily indoctrinated with colorblind ideology (racism). Rather, they think identifying as colorblind can be a sign an individual is interested in overall racial equity. Although this may be true, they warn identifying as colorblind often causes individuals to have negative reactions to, or perceptions of, policies and legislation designed to promote racial or economic equality (food stamps, affirmative action, discipline policies, etc.). Hartmann et al. suggested this paradox is a clear sign individuals are internalizing their views on race and culture. They believe individuals who strongly identify as colorblind no longer fit in colorblind ideological theory. The line of reasoning implies individuals fitting this description are emotionally fragile as they wrestle with their awareness of race and societal inequalities. Sue et al. (2007) report European Americans identifying as color blind often do so as a strategy to present themselves as non-prejudiced or mask prejudice in socially sensitive situations. Blitz et al. (2016) theorized the negative reaction individuals had concerning cultural responsiveness training in their study could have negative and unintended consequences. Teachers may become culturally blind; they might become resentful and unconsciously deliver microagressions and further alienate their African American students (Blitz et al., 2016). These findings suggest forced cultural responsiveness training is not a standalone (or effective) intervention able to reduce discipline disproportionality. When viewed in the context of Hartmann et al.'s study, however, a more individualized and self-directed intervention may produce different results.

Allen (2015) attempted to address disproportionality in a different manner in his qualitative study, utilizing critical race theory as the frame-work, with an anti-deficit lens.

He selected African American participants considered academically successful. Allen found these students were able to overcome racial barriers through resiliency, familial support, and intrinsic motivation to succeed later in life. He reported these particular students thought it ludicrous they would fail on purpose to avoid being called out as "acting white". Allen was highly critical of teachers–specifically teachers who had lower expectations for African American students. He reported his participants held special admiration for teachers who pushed them academically. While the results of this study are not generalizable to the African American population as a whole, his research does suggest discipline disproportionality is a classroom and teacher level problem. Allen's decision to employ critical race theory with an anti-deficit lens could indicate a paradigm shift and help guide how discipline disproportionality is studied and addressed.

Boneshefski and Runge (2014) noted discipline data is an essential and underutilized tool in the quest to reduce or eliminate discipline disproportionality. They asserted school level teams should review disaggregated discipline data in a consistent and habitual manner. The proliferation of powerful student information systems (SIS) and software designed to calculate risk among subgroups like School-Wide Information Systems (SWIS) make it easier for school and district leaders to monitor and address disproportionality during the school year (Boneshefski & Runge, 2014). They suggested this advance in data software could play a major role in schools hoping to reduce discipline disproportionality. They believe as teachers become familiar with this technology they should become cognizant of their discipline tendencies. Amin (2017) supported this hypothesis in a study on judge's courtroom decisions (verdicts, sentencing, etc.). Amin predicted if judge rulings and sentencing tendencies were scrutinized for

racial inconsistencies, judges would become cognizant of any decisions tainted by implicit bias. By facing the data, judges would be forced to recognize and account for their implicit bias–or face claims impartiality is devoid in their courtroom. Similarly, Gregory, Allen, Mikami, Hafen, and Pianta (2014) espoused the use of a professional development program designed to help teachers understand their role in reducing discipline disparity. The My Teaching Partner-Secondary (MTP-S) is an intensive 1-year program found effective at reducing the discipline gap between African American and European American students. A cornerstone of the program includes a systematic, yearlong review of teacher discipline decisions and interactions with African American students in particular. These three related studies suggest data review and subsequent awareness may aid in the effort to reduce discipline disproportionality.

Heilbrun, Cornell, and Lovegrove's (2015) study dealing with principal attitudes yielded some interesting and troubling findings. They were granted access to school safety audit surveys all high school principals in the State of Virginia were required to complete. The survey investigated whether principals thought zero tolerance policies were effective and necessary in their schools. They correlated principal responses with discipline data to ascertain if principals who thought zero tolerance policies were effective had higher overall discipline and discipline disproportionality rates. Heilbrun et al. (2015) found principals who thought zero tolerance policies were effective and necessary had higher rates of exclusionary discipline events. Furthermore, they found African American students were suspended at twice the rate of European American students and were often written up for disruptive (subjective) rather than violent or aggressive offenses in schools classified as pro zero tolerance. Heilbrun et al. did not

claim causality from the results of this study, but indicated further research is needed examining classroom (teacher) level differences and how these differences may affect discipline disproportionality.

Implicit Bias Development and Psychology

This section will define and explain the psychology of implicit bias, explore the research on how people develop implicit bias, and review the literature on how implicit bias may be a contributing factor on discipline disproportionality. According to Greenwald and Banaji (1995), implicit bias is a concept and theory in development for over half a century. Implicit bias is defined as the unconscious or automatic response or decisions people make in social situations when they do not have the necessary time or cognitive capacity to make a thoughtful, or explicit choice. They distinguish the difference between implicit attitudes and implicit bias. Implicit attitudes are loosely defined as a predisposition towards an event, person, or group. People are aware of their implicit attitudes, but may not let others know they possess them. They explain implicit bias, on the other hand, is thought to operate unconsciously and occur instantly (social cognition). Greenwald and Banaji claim many people are not aware of their implicit biases–and have no idea how bias may be affecting their decision making process in everyday social situations. Yoon's (2012) case study illustrated how lack of implicit bias awareness can contribute to negative outcomes (achievement and discipline) for African American students. The teachers (European American females) in Yoon's study were aware of racial disparities in education (discipline disproportionality and the achievement gap), and claimed to have knowledge of pedagogical strategies designed to overcome

these inequities. What these teachers lacked, Yoon believed, was an understanding of their own implicit bias and how it was effecting their everyday classroom decisions.

Greenwald et al. (2015) believe implicit bias harms the African American population in two distinct ways. First, they argue even seemingly minor effects of implicit bias in a particular population (i.e. police department) can negatively affect most African Americans who interact with the police department (population). Second, they contend small discriminatory acts rooted in implicit bias have a compounding effect on African American's as they are repeatedly and systematically exposed to these actions. These findings are particularly important for school settings where teachers and students interact multiple times on a daily basis.

Implicit bias is thought to develop early in childhood. Qian, Heyman, Quinn, Messi, Fu, and Lee (2016) found implicit bias was evident in children as young as three in both Chinese and Cameroonian participants. They theorized implicit bias may evolve over time–an important concept in this study. This theory was based on the results of the Cameroonian participants. Participants took a race IAT using Black and Chinese faces. The youngest participants appeared to harbor negative bias against Chinese faces while the adults did not. The researchers believe societal class and status alter a person's implicit racial biases as they develop and mature. They proposed racial or ethnic implicit bias develops early in a child's life but is not necessarily permanent. In a similar study, Setoh, Lee, Zhang, Qian, Quinn, Heyman, and Lee (2017) also found evidence of implicit bias in children three years old. Their research design tested the correlation of implicit and explicit bias using Chinese and Indian faces. Interestingly, they found no correlation of implicit and explicit attitudes. In contrast to adults, however, children displayed

evidence of both implicit and explicit bias indicating they told the truth on the explicit bias assessment. Setoh et al. theorized 3-year old's were not yet aware of socially desirable answers. These two studies seem to indicate implicit bias develops early in childhood, and is likely altered, over time, by culture and society therefore making change possible.

Kubota, Peiso, Marcum, and Cloutier (2017) conducted a study measuring whether individual's self-reported contact with African American's as a child and as an adult had an effect on IAT scores. They theorized implicit bias is shaped by stereotypes and prejudiced attitudes formed early in life. They reported both African American and European American participants displayed pro-white racial implicit bias. However, European American individuals reporting higher than normal childhood and current contact with African Americans had lower pro-white racial implicit bias. Interestingly, African Americans scores were only effected by current contact with African Americans– childhood contact was not found significant. They believe quality of contact, rather than quantity, played a role in these results. By utilizing a self-reported measure, they were able to illustrate the malleable nature of implicit bias.

Implicit bias is a well-established but poorly understood phenomenon (Devine et al., 2012; Greenwald & Banaji, 1995; Greenwald et al., 2003; Plant & Devine, 2009). More research is needed to understand the role implicit bias plays in the everyday decisions of people in general and teachers specifically. Glock and Klapproth (2017) found ethnic majority German teachers held implicit negative racial bias against minority Turkish students, regardless of age or gender. They did not explore how this bias may effect discipline decision making but did call for future research in this area. None of the

previous studies were based on the black/white dichotomy. It is important to establish implicit bias is likely shaped by cultural and societal expectations and experiences (Setoh et. al. 2017). For instance, Hannon, Defina, and Bruch (2013) found the skin tone of African American students seemed to make a difference in discipline disproportionality. They observed African American students with darker skin tone were more likely (than African American students with lighter skin tone) to be disciplined in disproportionate numbers–especially for African American females. Hannon et al. believe this was due to cultural stereotypes often depicting dark skinned African Americans females as aggressive and loud–both behaviors frowned upon in the educational setting.

Establishing race and skin tone as societal influences is an important part of understanding implicit bias (Dovidio & Gaertner, 2000). Eastman (2015) wrote a powerful qualitative analysis on race and culture. He analyzed over 8,000 anonymous newspaper comments written about two different motorcycle rallies near Myrtle Beach, SC. One rally was attended by predominantly affluent European Americans (Harley Davidson Motorcycle Rally) while the other was attended by mostly affluent African Americans (Memorial Day BikeFest) (Eastman, 2015). Eastman framed this study through the lenses of white innocence and black deviance, respectively. Both events would be described as scenes of debauchery. Although rally attendees were nearly identical in age, wealth, and educational level, and the crime rates similar for both events, the community opinion on the separate rallies were strikingly different. Harley Davidson rally attendees were mostly described as lawyers, veterans, and doctors enjoying some well-deserved time off. The African American rally attendees, on the other hand, were categorized as drug dealers, thugs, and rapists pillaging the town. With predictable

certainty, the community lamented the intense police presence at the Harley Davidson rally. A few weeks later, many writers requested the presence of the National Guard and mobile jails for the African American rally. In a similar study investigating race and venture capital, Younkin and Kuppuswamy (2017) found African American entrepreneurs received fewer and smaller contributions than European American entrepreneurs with similar backgrounds. Okonofua and Eberhardt (2015) found teachers were likely to recommend more punitive punishment and were more likely to escalate discipline procedures based simply on the name of a student (Jake vs. Jamaal). The body of literature exploring how differences in race shape American culture is beyond the scope of this study. The context these studies provide, however, are vital when discussing implicit bias and the possible effects this bias may have on discipline disproportionality.

Although the research on implicit racial bias is substantial and growing, it is not yet understood or accepted by society at large. Dovidio and Gaertner's (2000) study explained this anomaly using the aversive-racism framework. They theorized modern racial bias is conducted in a manner making it difficult to identify racially biased individuals. Explicit bias is generally frowned upon by society–especially among well-educated populations such as teachers (Dovidio & Gaertner, 2000). Smolkowski et al. (2016) believe most European American teachers are not overtly racist and go out of the way to appear non-biased or colorblind. They contend most teachers strive to appear non-discriminatory in decisions made deliberately. They further claim most teachers do harbor implicit bias against African American students and are not aware of it. They coined the term "vulnerable decision points (VDPs)" (Smolkowski et al., 2016) to

describe how these factors contribute to discipline disproportionality. They found distinct patterns in the discipline data suggesting specific conditions would trigger an ODR and if the offense was classified by the teacher as minor or major. McIntosh et al. (2017) also proposed most teachers are not overtly racist or aware of their implicit bias. In a previous study, McIntosh et al. (2014) dichotomized decision making at schools as system one and two. System one was thought to be efficient and automatic while system two decisions needed deliberate attention and required effort. They continue, by making teachers aware of their implicit bias (providing scenarios or vignettes illustrating how bias may effect discipline decisions), and giving teachers strategies to avoid making decisions based on bias, it would be possible to reduce their implicit bias (and discipline disproportionality).

In contrast, Fazio and Olson (2003) argue it is important not to associate implicit with unconscious. Just because a person displays implicit racial bias does not mean they are unaware of their bias. They speculated this may be the reason scores correlate differently on the explicit/implicit bias for race than for other social cognition measures (gender, politics, weight, etc.). This line of reasoning controversially suggested people who receive scores indicating implicit racial bias on the IAT probably were somewhat racist but were unwilling to admit this on their explicit assessment. Although the disagreement among the scholars is significant, Fazio and Olson's (2003) theoretical premise does conclude that, with effort, implicit bias can evolve over time. If a person wants to be or *seem* less racist they can lower their implicit bias score in repeated IAT measures. Both McIntosh et al. (2014) and Fazio and Olson (2003) come to the same

conclusion–making people aware of their implicit bias is the first step in helping people reduce implicit bias.

Implicit Association Test (IAT) and Reducing Implicit Bias

Project Implicit has a number of assessments designed to measure implicit bias on various social issues. The Race IAT, used in this study, is comprised of seven timed trials or modules. The first block is a training module and requires participants to categorize "good" or "bad" words (joyful and tragic) by pressing a pre-determined key. For good words a participant would strike the "E" key and bad words the "I" key. The second block is also a training module, and participants are asked to distinguish between African American and European American faces by pressing one of the pre-determined keys with automaticity (Greenwald et al., 2003). This is a time sensitive assessment and delays in latency are measured as the basis for determining the participant's implicit bias D-score. After the training blocks, participants are tasked with following the same procedure using the same keys to categorize positive and negative images and distinguish between African American and European American faces or vice versa. The next block switches instructions and asks participants to press either the African American key for positive images and the European American key for negative or vice versa. The next trial is the exact opposite as participants hit the European American key for positive and the African American key for negative (or vice versa) and so on and so forth. Once complete the assessment measures the different response times between trials. Faster response times for a combination (example European American/negative) would indicate the participant has implicit bias for that group (Greenwald et al., 2003). Greenwald et al. asserted the predictive validity of the race IAT has been stable over time.

Peterson et al. (2016) utilized a different IAT designed to measure implicit bias based on ethnicity in New Zealand. Their study measured the ethnic achievement gap between ethnic European and Asian students rather than discipline disproportionality. The results of the study indicated an intervention designed to reduce negative implicit bias was successful. Teachers seemingly reduced or compensated for their bias and apparently did close the achievement gap in reading for students in the study (Peterson et al., 2016). They called for more research on the effects of implicit bias across a litany of educational contexts.

Clark and Zygmunt (2014) propose most people will react in one of five different ways when confronted with results of the Race IAT (disregard, disbelief, acceptance, discomfort, and distress). The participants in their qualitative study were in-service teachers pursuing a graduate degree. Clark and Zygmunt were able to monitor participant reactions to IAT results by reading discussion board posts, a required component of the online class they were teaching. Teacher reactions and subsequent discourse gave Clark and Zygmunt a better understanding of how motivated teachers think through the implication implicit bias may negatively affect their pedagogy and overall effectiveness as an educator. These teachers, by in large, had reactions of discomfort or distress when faced with their results. On the other hand, teachers who reacted with disregard, disbelief, or acceptance were either unable or unwilling to grapple with the idea (implicit bias may reduce their effectiveness as an educator) (Clark & Zygmunt, 2014). Howell et al. (2014) warned future researchers to proceed with caution when utilizing the Race IAT. They found participants who took the IAT and had a large discrepancy between their explicit (self-reported) racial attitudes and implicit bias results were very defensive

when faced with the scores. They further asserted these defensive mechanisms can cause individuals to refute the results and become more entrenched in their racial beliefs. Greenwald et al. (2003) believe this is due to the social sensitivity of the subject (race). Interestingly, they found this is not true for other socially sensitive subjects. For instance, participants who took the IAT on politics were more likely to over report their political preference on the explicit assessment while their implicit political score was often considered moderate (Greenwald et al., 2003). They found the same participants were likely to do the exact opposite when faced with the Race IAT. Plant and Devine (2009) thought intrinsic motivation is largely responsible for this discrepancy. Their three studies recruited participants who indicated they were either internally or externally motivated to not appear prejudiced. They utilized the Attitudes Toward Blacks Scale (ATB) to measure this difference. An internally motivated (IMS) participant would likely respond positively to the statement "Being nonprejudiced is important to my self-concept" while an externally motivated (EMS) participant would select "I try to hide any negative thoughts about Black people in order to avoid negative reactions from others" (Plant & Devine, 2009, p. 642). The participants were offered a program purported to reduce the appearance of being prejudiced. Plant and Devine found in the second study high IMS participants did not think they needed the program because they felt they were free of prejudice while high EMS participants spent considerably more time and energy attempting to reduce their prejudice before taking the Race IAT. Even more, high IMS participants who got results indicating implicit racial bias would go back and spend time in the prejudice reducing program (Plant & Devine, 2009). The third study yielded somewhat controversial results. They found high EMS participants were very interested

in reducing socially detectable prejudice but were not very interested in reducing socially *undetectable* prejudice–i.e.–implicit bias. This led Devine et al. (2012) to develop an intervention designed to reduce implicit bias.

Plant and Devine (2009) believe reducing implicit bias would be possible if two conditions were met. "First, people must be *aware* of their biases and, second, they must be *concerned* about the consequences of their biases before they will be motivated to exert effort to eliminate them" (Devine et al., 2012, p. 3). Devine et al. (2012) conducted a study designed to measure the effectiveness of the aforementioned intervention at reducing implicit bias. Undergraduate psychology majors were asked to participate and were subsequently assigned to either the control group or intervention group. Both groups took the Race IAT and were given their results. The intervention group was then given strategies and guidance on how to overcome their bias while the control group was not. The authors found the intervention effective at reducing implicit bias *and* seemed to increase the groups concern for the consequences implicit bias can create. In fact, Devine et al. (2012) found the level of concern or awareness grew over time. This transformation of explicit views on race and culture among intervention group participants is a promising yet still speculative result of the study.

Pepis (2017) built on Devine et al.'s (2012) work by changing some of the conditions such as feedback and participant makeup while maintaining the use of the intervention and repeated measure IAT scores. The participants in Pepis' study were undergraduate pre-service elementary teachers. This shift in participant makeup is important as it relates to possible effects teacher implicit bias may have in school settings. She altered the groups so all participants received the intervention and only one

group received IAT feedback. The results of Pepis' were not as "robust" as those in Devine et al.'s, but promising nonetheless. She found all pre-service teachers had implicit biases on both race and socioeconomic status. Surprisingly, Pepis found participants in the group who did not receive IAT results achieved a greater reduction in their implicit bias score over time and were more likely to employ the strategies delivered in the intervention than the group that did receive their IAT results. She acknowledged this may be due to defensive reactions from the scores confirming Howell et al.'s (2014) warning.

Summary

This literature review covered a wide range of important topics in K-12 education and equity in school discipline. The body of literature on discipline disproportionality is substantial. There is strong consensus from both scholars and education officials discipline disproportionality is both pervasive and systemic (Losen et al., 2014; U.S. Department of Education, 2016). The literature on the effects of discipline disproportionality on African American students is both vast and saddening. The number of theories and possible solutions are considerable yet no current practice on this deleterious phenomenon has been able to end the inequity. Even worse, the problem seems to be growing despite these efforts. This study fills a gap in the literature concerning the possible interaction implicit bias may have on equity in school discipline. Research on implicit bias goes back nearly 30 years and covers a wide range of social constructs including poverty, gender, and race to name a few. Race based implicit bias is considered socially sensitive and is only now being considered for K-12 research. Studies in other regions and nations have provided both troubling and promising results.

Plant and Devine (2009) think individuals must be aware of their implicit bias and be concerned or vested in the possible consequences of implicit bias before they can change. By recruiting in-service teachers concerned about the topic of equity in school discipline this study tested their theory. This study provides context and data in the ongoing efforts to reduce or eliminate discipline disproportionality.

Chapter III

METHODOLOGY

This chapter explains the methods used in this study. Many of the methods chosen are similar to research previously conducted by Devine et al. (2012), Forscher et al. (2017) and Pepis (2017). The minor variations allowed the researcher to build on their work. The first section contains a description of the research design and why the design was chosen. Section 2 describes the population utilized in this study. The next section describes the intervention, instruments, and data collection used in the study, and the fourth section describes how and when both quantitative and qualitative data was gathered. The fifth section contains the quantitative data analysis, along with statistical considerations and assumptions, and qualitative data analysis. A chapter summary briefly recounts the methodology used for this study.

Research Design

This study employed the explanatory sequential research design. This mixed methods approach uses both quantitative and qualitative procedures in sequence. Quantitative data was first gathered and analyzed. The quantitative results were then used to guide the qualitative phase. Participants for the qualitative portion were purposefully selected to explain the results and group differences or characteristics discovered in the quantitative phase. This design was chosen due to limited population, time, and funding concerns. Creswell and Plano Clark (2017) used the word *pragmatism* when describing mixed methods research. Mixed method designs are often more

concerned with understanding the problem at hand (research question) rather than methods or procedures. Creswell and Plano Clark further characterize mixed methods design as oriented in practicality rather than theory–in other words research designed to understand "'what works' [in] real world practice" (p. 37). Creswell and Plano Clark explain this design is well suited for research attempting to help in the development of approaches still evolving. The qualitative data was useful in determining if teachers thought implicit bias awareness and concern effect discipline patterns in their school or classroom. By using the explanatory sequential design, it was possible to understand the types of motivation prompting teachers to volunteer for the study and if different motivators led to participants reducing their implicit bias (or not).

Quantitative data was first gathered and analyzed. The analysis of the quantitative data informed what questions (qualitative) needed context or clarification and identified differences in participant perception of the Implicit Association Test (IAT) feedback (D-score) and Devine's intervention. The independent variables in this study were repeated measure IAT assessments and Teacher Multicultural Attitude Survey (TMAS) scores. The dependent variables in this study are (a) bias as measured by the IAT at stage 1, 2, and 3, and (b) differences in awareness and sensitivity of multicultural issues in the classroom as measured by the TMAS. The TMAS and IAT scores are measured at the interval level.

The qualitative portion of the study identified strengths and weaknesses in this study based on feedback generated by volunteer, in-service teachers. Maxwell (2013) explained this process allows the researcher to craft interview questions both specific (past-tense) and generalized (present-tense). Interview questions designed in this

manner, Maxwell continues, allow participants an opportunity to clarify how a variable or outcome was effected due to changing perspectives. The analysis of the qualitative data framed how future research on equity in school discipline with comparable populations could be improved and provide guidance on if taking the race IAT is recommended for individuals interested in reducing their implicit bias.

Participants

Participants were volunteer teachers from a rural school district in the southeastern United States. The quantitative portion of this study took place during April and May of 2020 and the qualitative portion in August 2020, both during the height of the COVID-19 pandemic. The district serves a population whose residents, on average, earn less than the state or national average. District data indicated nearly 70% of certified employees had advanced degrees while the community the district serves had a rate nearly half the degree rate (bachelor's degree or higher) of the state average. District student makeup contained a disproportionate number of students in poverty, dual language learners, students effected by childhood trauma and toxic stress, and migrant students than the average in the state and nation. The participants all taught at a school in the school system, (grades pre-K through 12th grade), that were accessible through approval of district leadership. Although Onwuegbuzie and Leech (2004) recommend over 201 participants for studies designed in this manner, the more realistic goal was the 21 participants per group deemed acceptable in experimental studies by Onwuegbuzie and Collins (2007). A total of 678 in-service teachers were recruited via email to participate in this study. The initial email (appendix A) asked teachers if they were "concerned about the topic of *equity in school discipline?"*. District data indicated over

800 certified staff work in the school system, but only in-service teachers (678) were invited to participate. Participants were offered three small tokens of appreciation for participating in the study. The tokens included a "jean pass", "early-leave pass", and a mini blizzard from Dairy Queen. The tokens of appreciation were delivered to the schools at the beginning of the next school year. From the initial email, 91 teachers (13.4%) responded indicating they were concerned about equity in school discipline and agreed to participate. All email addresses were entered into a Qualtrics contact list for survey distribution. Individual links were emailed to all participants. Due to the mandatory school closures in place, participants all took part in the study at home. All responses were anonymized. Of the 91 links sent, 13 did not complete the survey for time point 1. A total of $N = 78$ participants completed the TMAS at time point 1 for this study (see Figure 1 for participant flow). A total of $N = 60$ participants completed the posttest TMAS and were included in the quantitative statistical analysis for research question 1. Basic demographic information was gathered during time point 1 (see Table 1). Other demographic information included years of experience, education level, and their current grade level teaching assignment (see Table 2). According to the National Center for Education Statistics, public school teachers in the United States were 76 percent female and 79 percent White (U.S. Department of Education, 2020). The participant demographics who completed time point one skewed from the national average in both gender (92% female) and race (91% White). This difference, coupled

with the small sample size and participant attrition, greatly limits the generalizability of the results from this study.

Participants were randomly assigned to the control (n = 39) or experimental (n = 39) group after taking the TMAS. Of the 39 participants in the experimental group, only 26 completed the IAT (IAT1). The control group had five participants drop out of the study at time point 2. The experimental group had 15 participants drop out at time point 2. Only 15 of the 24 participants for time point 2 took the IAT (IAT2). The control group had 34 participants complete the posttest TMAS. The experimental group had 26 participants complete the posttest TMAS. Only seven participants completed the time point 3 IAT (IAT3).

Table 1

Demographic Characteristics of Participants at Baseline

Baseline characteristic	Control Group		Experimental Group	
	n = 39	%	n = 39	%
Gender				
Female	36	92	36	92
Male	3	8	3	8
Age				
21-24	0	0	0	0
25-34	10	26	9	23
35-44	12	31	15	38
45-54	14	36	14	36
55+	3	8	1	3
Ethnicity				
Hispanic	0	0	0	0
No	39	100	39	100
Race				
White	35	90	35	90
Black or African American	4	10	35	3
American Indian or Alaska Native	0	0	0	0
Some other race	0	0	0	0

Table 2

Professional Characteristics of Participants at Baseline

Baseline characteristic	Control Group		Experimental Group	
	n = 39	*%*	*n* = 39	*%*
Experience (years teaching)				
0-3	1	3	2	5
4-7	6	15	10	26
8-12	13	33	5	13
13-18	9	23	12	31
19+	10	26	10	26
Education (highest degree attained)				
Bachelor's	9	23	8	21
Master's	21	54	16	43
Specialist's	9	23	15	38
Doctorate	0	0	0	0
Current teaching assignment				
Elementary (pre-K – 5)	23	59	28	72
Middle (6 – 7)	2	5	6	15
Junior High (8 – 9)	7	18	3	8
High School (9 – 12)	7	18	2	5

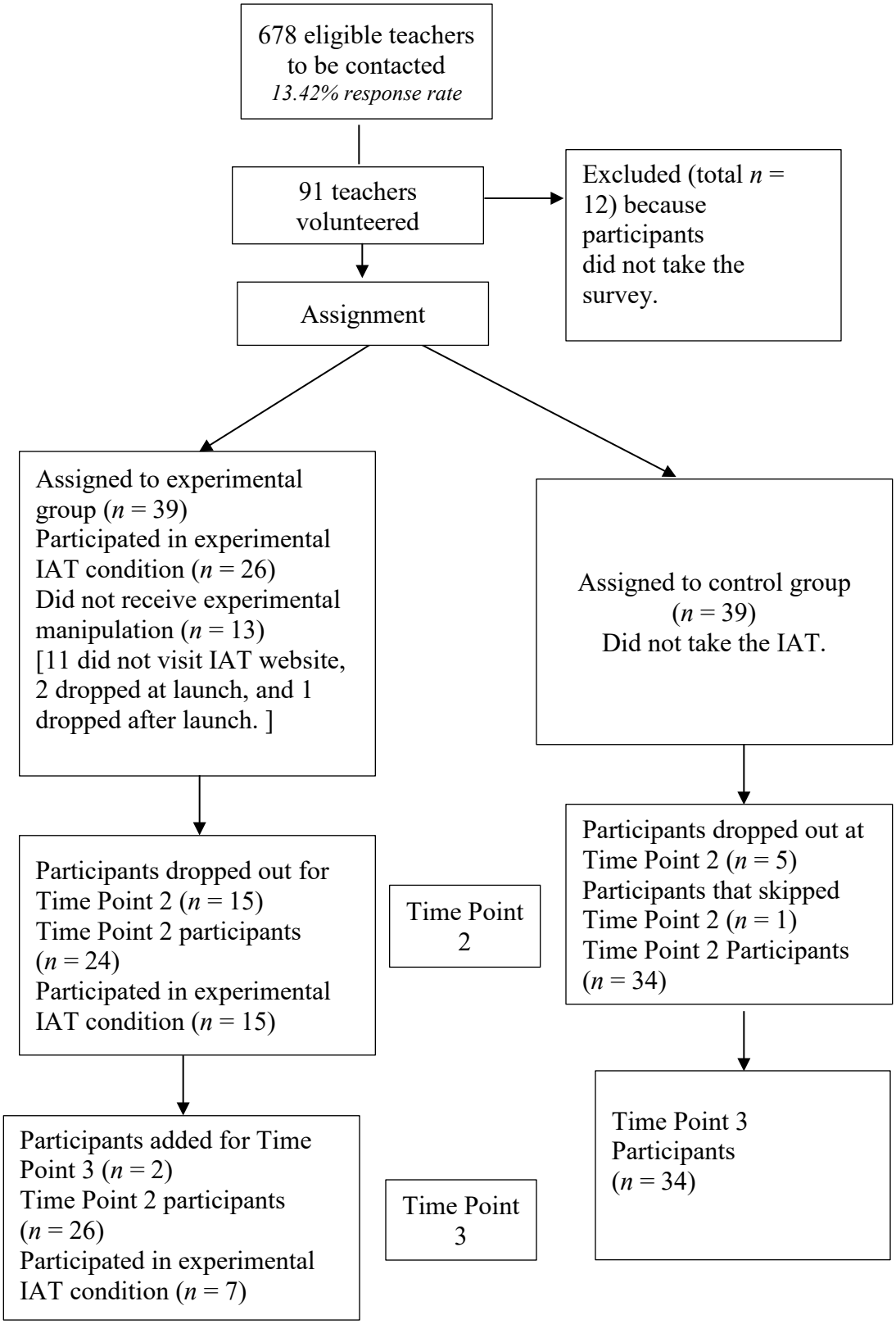


Figure 1. Flow of participants for the given study procedures, assignment (time point 1), time point 2, and time point 3.

Ary, Jacobs, Sorenson, and Walker (2014) suggest that volunteer sampling is a subset of convenience sampling–and convenience sampling is nearly always considered the weakest type of nonprobability sampling. Creswell and Plano Clark (2017) suggest with correct planning and quantitative data analysis, this type of sampling will yield specific, detailed, results. Specific differences in participant characteristics create an opportunity for qualitative data likely to provide in-depth explanation and understanding of participant perception. Maxwell (2013) stated structuring qualitative inquiry in this manner is highly effective when comparing differences in participants.

Participants for the qualitative portion of this study were purposefully selected. All participants were asked to participate in the qualitative portion of the study. Participants who indicated they were willing to be interviewed were asked to take part in an interview. Thirteen teacher volunteers representing elementary, middle, and high school teachers were identified. Consent was obtained from the participant and the principal of the school where they work. Interviews were semi-structured to gain context or deeper understanding on why they thought participating in the study was useful (or not).

Instrumentation

Quantitative Component

Two instruments were utilized in the quantitative portion of this study. The Teacher Multicultural Attitude Survey (TMAS) was administered to all participants as a pretest and posttest measure. This instrument is designed specifically for teachers. Scores indicate if a teacher thinks multicultural issues, as it relates to teaching and learning, are important (or not). The race Implicit Association Test (IAT) was then

administered to the experimental group. IAT results measure participants bias towards Black or White skin tone, respectively.

The Teacher Multicultural Attitude Survey

The Teacher Multicultural Attitude Survey (Ponterotto, 1995a) was administered to participants through the Qualtrics platform at the beginning of stage 1. The TMAS is a brief (20 question) survey designed for teachers. The survey is widely considered non-threatening (non-evaluative) and is by in large accepted as a valid psychological instrument (Ponterotto, Baluch, Greig, & Rivera, 1998). Participants rate each of the 20 questions on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Seven of the items (3, 6, 12, 15, 16, 19, 20) are reverse-scored. Scores can range from 20 to 100. A score of 20 would indicate less interest in multicultural issues affecting classrooms while a score of 100 would indicate great interest. All scores were confidential and only matched to the login ID created by Qualtrics.

Validity. Establishing an instrument is valid and essential when attempting to measure psychological constructs–in this case, awareness and sensitivity of multicultural issues in the classroom (Ary et al., 2014). The TMAS was developed by a team of four individuals; a university professor and three graduate students. Test items were evaluated and altered to ensure the single-factor modeled instrument measured "teachers' sensitivity and appreciation of cultural diversity" (Ponterotto et al., 1998, p. 1006). The original instrument was revised down to 20 questions and once again evaluated for construct validity. The questions were constructed with either negative or positive wording (7 negative, 13 positive) and has no subscales. The revised TMAS closely correlated with comparison instruments (convergent correlations) and deemed a valid

instrument for measuring teacher awareness of cultural issues, such as racial bias, in the classroom and school (Ponterotto, et al., 1998). The comparison instruments used to establish convergent validity were Ponterotto's (1995b) Quick Discrimination Index (QDI) racial and gender equity subscales and Phinney's (1992) Multigroup Ethnic Identity Measure, Other Group Orientation (MEIM-OGO) subscale. Correlation coefficients reached moderate levels between TMAS and QDI racial subscale (.45), QDI gender equity subscale (.35), and MEIM-OGO subscale (.31). These correlation scores, although not very high, did provide adequate evidence to claim convergent validity for the TMAS. Ponterotto et al. (1998) also utilized the group differences approach by analyzing differences in scores among naturally occurring groups in the sample ($N = 227$) to establish criterion-related validity of the TMAS. Ponterotto used three *t*-tests to compare scores for European American and non-European American's, females v. males, and participants who had some training in multicultural awareness and those who had not received prior training on multicultural awareness. Ponterotto conducted a fourth univariate test using a one-way analysis of variance (ANOVA) to compare groups by the amount of multicultural training they had received. Four groups were created based on training level. One group had no multicultural training. A second group indicated they had never attended a formal training but had been exposed to multicultural awareness in other trainings or courses. A third group reported only one formal training course in multicultural awareness. The fourth group reported attending two or more formal training sessions on multicultural awareness. After reducing alpha using the standard Bonferroni formula to .0125 to account for four univariate tests (.05/4), only one *t*-test produced significant results. Participants who had some training ($n = 86$, $M = 85.3$, $SD =$

7.7) scored higher than those who had not attended multicultural training (n = 136, M = 80.0, SD = 8.9) producing a significantly different result, $t(198) = 4.68$, $p < .001$.

Reliability. Ary et al. (2014) categorize reliability as a measure of consistency– will the instrument generate similar results in comparable populations given at different times. The TMAS was analyzed for reliability in two separate ways. Ponterotto et al. (1998) utilized test-retest stability assessment and two measures of internal consistency. The TMAS was administered to a group of 16 graduate students over a three-week period to establish test-retest reliability and stability, producing an alpha of .80 (Ponterotto, et al., 1998). Internal consistency measures produced a Cronbach's alpha of .86 and a theta coefficient of .89. Cronbach's alpha is an acceptable method of establishing internal consistency for instruments containing Likert scale items. For standardized items, like those found in the TMAS, the alpha coefficient is calculated using the coefficient of variation. The TMAS coefficient of variation was 9.3%. Theta coefficient, on the other hand, is dependent on a principal component analysis. Ponterotto et al. reported the TMAS principal component analysis produced results indicating the instrument is a reliable instrument for the single factor of multicultural sensitivity and awareness with structure coefficients above .35 for 18 of 20 items. These scores were within acceptable limits to establish the internal consistency of the TMAS and claim the instrument reliable.

The Implicit Association Test

The Race IAT is an instrument first released by Greenwald, McGhee, and Schwartz as the Implicit Association Test (1998). The IAT was designed to assess individual bias through association and response time (latency). The IAT contains 7 sections, or blocks, where participants quickly categorize a picture, symbol, word, or

combination of the two, by pressing a specific key on the left or right of the computer keyboard. Blocks 1, 2, and 3 contain 20 items (or trials), blocks 4 and 5 contain 40 items, block 6 contains 20 items, and block 7 contains 40 items. At the beginning of each section of the Race IAT, participants are given instructions on what key to press, right or left, when they see a particular image (black face or white face) and word category (good/bad). Nosek, Smyth, Hansen, Devos, Lindner, Ranganath, Smith, Olson, Chugh, Greenwald, and Banaji (2007) provided a clear example using the social construct of gender. In section 1, participants are instructed to hit left key (alphanumeric key "E") for pictures of female faces and right key (alphanumeric key "I") for pictures of male faces. Block 2 instructs participants to hit left key for pleasant words and right key for unpleasant words. Blocks 3 and 4 use left key for female faces and pleasant words and right key for male faces and unpleasant words. Block 5 uses left key for male pictures and right key for female pictures. Blocks 6 and 7 use left key for male pictures and pleasant words and right key for female pictures and unpleasant words. The entire assessment takes less than 10 minutes to complete. In this example, participants who favor female construct would have faster response times and few errors when women are associated with good words in blocks 3 and 4, while participants who had a preference for male construct in blocks 6 and 7. Participant scores are computed and assigned a D-score (Greenwald et al., 2003). Scores range between -2 to +2. Scores above zero in the above example would indicate a preference for female construct and scores below zero indicate preference for male construct. The most recent copyrighted version of the assessment in this study is known as the Race IAT. Race IAT D-scores were presented to participants once they finished the assessment. A D-score less than 0.15 means an

individual has little to no preference for white skin. A D-score between 0.15 and 0.35 indicates an individual has a slight preference for white skin. A D-score between 0.35 and 0.64 means an individual has a moderate preference for white skin and a D-score more than 0.64 would mean an individual has a strong preference for white skin. The same scores and descriptions hold true for negative D-scores and black skin. Over time, Project Implicit (2011) has developed numerous tests assessing individual bias in many different areas (disability, weight, age, etc.). Both Devine et al. (2012) and Pepis (2017) utilized the Race IAT to assess an individual's level of implicit bias over time after taking an intervention designed to help individuals reduce their bias (if applicable) against African Americans.

Validity. The Race IAT assesses bias through association and latency. Instrument validity is well established–it has been tested in scores of studies since it was released in 1998 (Greenwald et al., 2003). A large-scale analysis spanning 6 years with over 2.5 million IAT assessments was conducted by Nosek et al. in 2007. The Race IAT was taken 732,881 times (Nosek et al., 2007). They reported both predictive and construct validity have been stable over time. In some cases, particularly when studying socially sensitive subjects (race), the predictive validity of the IAT was higher (.25) than self-reported measures (.13). In studies measuring other social constructs, like political or brand preference, explicit measures were better predictors (.71) than the IAT (.40). Nosek, Greenwald, and Banaji (2006) reported the IAT had sufficient predictive validity for use in educational research. Construct validity scores on latency-based measures are often far lower than other psychological assessments. In a more recent publication, Greenwald et al. (2015) analyzed a meta-analysis questioning the predictive validity of

the Race IAT (Oswald, et al. 2013). Greenwald et al. (2015) characterized the IAT as having strong "*consequential validity"*, which is considered part of construct validity (p. 557). This type of construct validity is particularly useful when measuring socially sensitive constructs like implicit racial bias. Greenwald et al. (2015) reported small effect sizes have a significant impact on individuals in two distinct ways: (a) one person acting in a discriminatory manner can affect many different people and (b) a person who experiences bias can experience the negative event many times–creating a powerful cumulative effect. Consequential validity is extremely important when trying to assess traits many would consider socially undesirable. Simply taking the assessment, Greenwald et al. theorized, can produce a small effect size. A small effect size viewed with a cumulative lens can produce a significant systemic impact, a concept Merton (1968) coined the *Matthew effect.*

Reliability. The IAT is also considered a reliable instrument (Greenwald et al., 2003). Nosek et al.'s (2006) meta-analysis asserted IAT scores have adequate test-retest reliability (.69), far higher than other latency measures (.25). Greenwald and Nosek (2001) reported internal consistency measures remain stable over time and estimate coefficient alpha values typically range from .7 and .9. Nosek et al. (2007) analyzed over 700,000 race IAT scores and reported a coefficient alpha of .75, far higher than scores ranging from -.05 to .28 for other latency based assessments. Rezaei (2011) found IAT reliability increased as participants became more familiar with the instrument by taking the test on multiple occasions. The test-retest reliability index score was much higher (.57, .57, .56) for groups who were able to practice using the instrument before taking the assessment than for the group that did not (.32). Rezaei also reported participants

considered the IAT more reliable as they became familiar with the instrument. It remains unclear if this phenomenon is due to participant increase in awareness of implicit bias, reduction of implicit bias, or familiarity with the instrument.

Qualitative Component

Creswell and Plano Clark (2017) indicate mixed methods research should follow a set of procedures dependent on the design chosen for the study. When using the explanatory sequential design, the quantitative results must be checked for accuracy prior to establishing the questions in the qualitative portion. This ensures the qualitative question(s) are written to drill down and describe significant or insignificant statistical findings. Creswell and Plano Clark (2017) highlighted 3 inherent validity threats one must address when using the explanatory sequential design. To ensure qualitative validity, the quantitative data and results should be checked for accuracy and described sufficiently to participants to capture their thoughts on any surprising findings. The questions were given to the participants prior to the interview. Providing participants the questions prior to the interview helped solidify rapport between the interviewer and interviewee. Seidman (2012) indicated when exploring socially sensitive topics via interview, especially race, establishing strong rapport is a vital methodological component. Written consent was obtained from the principal of the school where each selected participant is employed.

Qualitative questions were initially based on Pepis (2017) interview questions and reviewed by an expert panel (see Appendices B and C). All panel members had doctorates in education and provided feedback on the presentation of quantitative results and question order and syntax. The questions were revised (see Appendix D). The final

questions (Appendix E) were established after the first two participants suggested adding a timeline of the study, definition of key terms, and a question concerning the role of K-12 leadership when addressing equity in school discipline. Codes were created based on anticipated themes garnered from initial interviewee data. Accuracy was established by recording and transcribing all interviews.

Data Collection

Quantitative Data Collection

Once the Institutional Review Board granted permission (see Appendix G), data was collected in the following manner. The race IAT was taken 3 times in stages–stage 2 followed stage 1 by one week, and stage 3 followed stage 2 by one week. Volunteer teachers (criterion) comprised the purposeful sampling group (participants), all received Devine's 2016 intervention and took Ponterotto's (1995a) TMAS.

Stage 1, TMAS was given initially to all participants ($N = 78$) and demographic information was collected. Demographic information included gender, age, education level, ethnicity, number of years in education, and current teaching level (elementary, middle, high). IAT (experimental) group participants ($n = 39$) were randomly assigned by Qualtrics. The participants were asked to take the IAT and complete a modified version of Devine's (2016) new intervention. Of the 39 participants in the experimental group, 12 did not complete the IAT. Control group participants ($n = 39$) did not take the IAT. They received a brief explanation of the IAT assessment and then completed Devine's intervention. Patricia Devine gave permission to use the updated version of her original 2012 intervention. The updated intervention includes a section where participants are asked to write an essay on intervention effectiveness–a design feature

added to help participants retain and utilize the strategies on overcoming racial bias embedded in the intervention. Devine's 2016 intervention was selected because it has been altered to increase effectiveness over the 2012 intervention–considered effective at reducing implicit bias as measured by the race IAT over time. Using volunteer participants extended her research by investigating if intrinsic motivation concerning implicit bias awareness improves intervention efficacy at reducing implicit bias over time. The intervention was embedded on the Qualtrics platform–it is an interactive, informative, educational presentation on what implicit bias is, and provides specific, pragmatic strategies proven to help motivated individuals reduce levels of implicit bias. Strategies include stereotype replacement, thinking of counter-stereotypic examples, individuating instead of generalizing, perspective taking, and increasing opportunities for contact (Devine, 2016). Once IAT participants finished the assessment, they were given their race IAT score and provided an explanation of what the score indicated embedded in the intervention. Forscher et al. (2017) posited timing feedback is critical–participants need to understand what the race IAT measures and how, so they understand how the test works (without becoming defensive).

Stage 2 occurred the following week. Experimental group participants logged into Qualtrics and were administered a second administration of the IAT. Of the 24 participants who logged into Qualtrics, only 15 took IAT2. Experimental group attrition is a concern during longitudinal studies and often threatens internal validity. Participants in the control group logged into Qualtrics and answered a question about COVID-19 and were excused. One week later, stage 3 included logging into Qualtrics and a third administration of the IAT for the experimental group. All participants then took the

posttest TMAS. Participants were asked to indicate if they found the study useful and/or effective. Stage 4 included interviews with volunteer participants. Interview questions were initially based on Pepis' (2017) study conducted with pre-service teachers (see appendix C) but were substantially altered.

Qualitative Data Collection

Qualitative participants were purposefully invited in an effort to cover the "heterogeneity in the population" (Maxwell, 2013, p. 98). By using this method, Maxwell (2013) explains, participants with different viewpoints or experiences will best explain the various attitudes participants develop over the course of the study concerning implicit bias and discipline disproportionality. Participant differences (age, experience, gender, current teaching assignment, school demographics) were also considered when recruiting them for the qualitative portion of the study. Participants were able to review their transcript for accuracy and clarify or expand on any answers they deemed inaccurate or incomplete. Confidentiality was ensured by using codes for participants. All qualitative data will be stored securely and disposed of as required by IRB policy.

Data Analysis

Quantitative Data Analysis

A one-way ANCOVA was used to answer research question 1. Data were first screened for missing data and outliers. Descriptive statistics were summarized and included the sample size for each group, as well as demographic information. The group means, skewness and kurtosis, standard deviation, and the standard error for of the difference between group means on the TMAS were calculated. The dependent variable (TMAS score) were continuous and measured at the interval level. Groups were

independent of each other and assigned randomly to either the control or experimental group.

A one-way analysis of covariance (ANCOVA) was used to determine if there was a significant difference in control and experimental group mean posttest TMAS scores after controlling for group mean pretest TMAS scores. The data was analyzed with plots to check linearity assumption between group pretest and posttest scores. The homogeneity of regression slopes was calculated and was non-significant. Normality of residuals assumption was checked by the Shapiro-Wilk test and homogeneity of variances assumption by Levene's test. Residuals were analyzed to check for outliers. Once all assumptions were met, output of the ANCOVA was reported to determine if there was a significant difference between the control group and experimental group posttest TMAS score means after controlling for pretest TMAS scores. Group means and standard deviation are reported. The *F* statistic was calculated at alpha level of .05, degrees of freedom, *p* value, and effect size (generalized eta squared) were reported from the results of the ANCOVA. Effect size was reported to discover practical differences between the groups.

A one-way analysis of variance (ANOVA) was used to answer research question 2. Research question 2 determined if participants who scored low or high on the pretest TMAS had significant differences on the final IAT (IAT3) scores. Data were screened for missing data and outliers. Descriptive statistics including the mean and sample size for each of the groups are presented. The standard deviation, group mean differences, and skewness and kurtosis were also reported and visually inspected. Prior to using an ANOVA, a Shapiro-Wilk test and was conducted and both histograms and Q-Q plots

created to check normality assumption. Groups were created with pretest TMAS scores within the normal distribution of scores found in the descriptive summary. Levene's test checked homogeneity of variance and box plots were visually inspected. Independence of observations assumption is assumed to be met as data does not include paired or repeated measures. A one-way analysis of variance was appropriate for question 2 because it had two independent groups on a single dependent variable and both dependent and independent variables were measured at the interval level. Once assumptions were met, output of the ANOVA was reported to determine if there were any significant differences between groups who scored low or high on posttest IAT score means. Group means and standard deviation are reported. The *F* statistic was calculated at alpha level of .05, degrees of freedom, *p* value, and effect size (generalized eta squared) are reported from the results of the ANOVA. Effect size calculated with Cohen's *d* is reported to discover practical differences between the groups.

A one-way repeated-measures ANOVA was calculated for research question 3. A repeated-measures ANOVA analysis measured participant repeated IAT scores over a 3-week period. Participant IAT scores were coded as time point 1, time point 2, and time point 3. Data was screened for errors, missing data, or outliers. A QQ plot and histogram was visually examined and Shapiro-Wilk normality test was checked to meet normality assumption. Group means, standard deviation, and the group mean differences were reported as descriptive statistics in a table. Mauchly's test was conducted for sphericity assumption. Mauchly's test was not significant and homoscedasticity assumptions were met. A repeated measures ANOVA was appropriate for this analysis since the independent variable had three or more levels on a single dependent variable

(IAT scores) and both dependent and independent variable were measured at the interval level. A repeated measures ANOVA was calculated and the *F* statistic with alpha level of .05, degrees of freedom, *p* value, and effect size of generalized eta squared reported to show any significant or practical differences on IAT repeated measure scores. If any significant differences between IAT scores had occurred, post-hoc comparison tests utilizing Bonferroni multiple testing method would have been conducted. The graphic output of an interaction plot was visually inspected to explore the effect of time on IAT score means and Cohen's *d* was reported for effect size between groups.

Qualitative Data Analysis

For the qualitative data analysis all interviews were recorded and transcribed. Teachers were invited to participate in the qualitative portion of the study at the end of time point 3. Teachers were purposefully recruited to capture their perception of study effectiveness. A total of 10 teachers indicated they would participate in the interview phase of the study. These participants were selected to examine similarities and differences in teacher views on multicultural issues in the classroom (Maxwell, 2013). Teacher views about multicultural awareness and sensitivity were monitored continuously to identify emerging areas of concern, themes in participant response, or categories so data could be organized efficiently. Qualitative data was also compared to quantitative results to see if teacher explicit views on multicultural issues supported quantitative data. Other potential qualitative questions, for example, could have centered on how the intervention could be integrated into current educator professional development requirements and if the intervention had a personal impact on specific teachers and why. Open-ended and follow up questions were noted if participants

indicated they had something different or important to share. Once interview data was transcribed, interview data was read and coded to determine themes emergent from the data. Participants had the chance to clarify any of their responses. Once complete, themes were used to add context and clarity to quantitative results.

Summary

This chapter served as a guide to the methodology of this study. The procedures in this chapter were designed to ascertain if volunteer, in-service teachers view multicultural awareness and sensitivity as an important issue. Once IRB approval was obtained, this chapter explained why a mixed methods explanatory sequential research design was chosen. A description of the population of study participants and a description of why this population was important, and how data was collected and analyzed for both quantitative and qualitative portions of the study were also included. Quantitative results indicate whether in-service teacher's sensitivity to multicultural issues change over time after being exposed to information on the topic. Quantitative results also measured if taking the IAT helped reduce the level of implicit bias as measured by the IAT over time, and how taking the IAT increased or decreased sensitivity to multicultural awareness. The qualitative results help explain the quantitative results and investigated discrepancies realized between teacher explicit (self-reported) and implicit (IAT results) views concerning multicultural awareness in the classroom. The intent of this study and the results garnered may help frame future research and policy aimed at understanding and eventually eliminating discipline disproportionality in our nation's schools.

Chapter IV

RESULTS

There were two purposes in conducting this mixed methods study. The primary purpose was to determine if there was a significant difference in participant's level of implicit racial attitudes by group (i.e., control and experimental) after completing an intervention purposefully designed to help individuals reduce or eliminate their bias. All participants completed the TMAS as a pretest and posttest measure. Control group participants did not take the IAT while participants in the experimental group did complete the IAT. Experimental group participants took the TMAS before the IAT at time point 1, took the IAT again at time point 2, and were asked to complete a third IAT at time point 3 before they took the posttest TMAS. The secondary purpose was to better understand participant's perception of implicit racial bias, the IAT, the effects of implicit bias on equity in school discipline, and how their views on these topics changed over time.

The following questions were answered in this study.

1. Is there a significant difference between pretest and posttest scores on Ponterotto's (1995a) Teacher Multicultural Attitude Survey (TMAS) by the control group and experimental group?
2. Is there a significant difference among participants who score low or high on the pretest TMAS on the final IAT score?

3. Is there a significant difference in participant repeated measure scores on the Implicit Association Test (IAT)?
4. In what ways do the interview data of licensed teachers about their views on the importance of awareness of implicit bias in classrooms provide an explanation for any quantitative results from the IAT or TMAS?

This chapter presents the findings of this study. The first section contains the descriptive statistics for TMAS and IAT scores and the demographic characteristics of the participants in this study. Participants who dropped out of the study were not included in the results of the quantitative findings of this study. The second section will report the results of the one-way analysis of covariance (ANCOVA) for research question 1, the results of a one-way analysis of variance (ANOVA) for question 2, and the results of a one-way within-subject repeated measures ANOVA used for question 3. The quantitative results guided the development of interview questions used for the qualitative component of this study. The third section includes the demographic characteristics of the interview participants and the fourth section includes the response data gathered through the interviews from both control and experimental group participants.

Demographic Characteristics of Participants

Participants were recruited via email to participate in the study. Email respondents were sent a secure link to the baseline survey and TMAS via Qualtrics software. Participants were given a secure login id and randomly assigned utilizing Qualtrics software to the control group (n = 39*)* or experimental group (n = 39) (see Appendix F for all mean pretest TMAS scores). Of the 78 teachers who started the study,

n = 34 from the control group and n = 7 from the experimental group completed the study. While only seven participants from the experimental group completed all three administrations of the IAT, 26 experimental group participants were exposed to the IAT and completed the posttest TMAS. The decision was made to include the 26 experimental group participants posttest TMAS scores in the one-way ANCOVA used to answer research question 1. Basic demographic information (see Table 3) was gathered and teachers indicated if they were concerned professionally about equity in school discipline, if equity in school discipline affected the climate of their school, and if equity in school discipline affected the academic performance of their students (see Tables 4 and 5).

Table 3

Demographic Characteristics of Participants who Completed the Pretest and Posttest TMAS

Baseline characteristic	Control Group		Experimental Group	
	n = 34	*%*	*n* = 26	*%*
Gender				
Female	32	94	25	96
Male	2	6	1	4
Age				
25-34	9	27	8	31
35-44	10	30	11	42
45-54	12	35	7	27
55+	3	9	0	0
Race				
White	32	94	23	89
Black or African American	2	6	3	12
Experience (years teaching)				
0-3	0	0	1	4
4-7	6	17	7	27
8-12	11	32	3	12
13-18	8	24	7	27
19+	9	27	8	31
Education (highest degree attained)				
Bachelor's	8	24	5	19
Master's	19	56	10	39
Specialist's	7	21	11	42
Current teaching assignment				
Elementary (pre-K – 5)	19	56	18	69
Middle (6 – 7)	1	3	5	19
Junior High (8 – 9)	7	21	1	4
High School (9 – 12)	7	21	2	8

Both the control group and the experimental group (at baseline) had 36 female and three male participants (see Table 1). Both groups also had 35 participants who reported their race as White. The control group had four participants who reported their race as Black or African American (see Table 2 for professional demographics of all participants). The experimental group had three participants who reported their race as Black or African American (one participant did not report their race). All participants in

the control and experimental group at baseline completed the pretest TMAS. An item analysis was calculated for the pretest and posttest TMAS (see Tables 6 and 7).

Table 4

Characteristics of Participants Before TMAS Pretest Administration

Baseline characteristic	Control Group		Experimental Group	
	n = 34	%	*n* = 26	%
Concern				
Yes	31	91	21	81
No	3	9	5	19
Academic				
Yes	29	85	21	81
No	5	15	5	19
Climate				
Yes	28	82	20	77
No	6	18	6	23

Table 5

Characteristics of Participants After TMAS Posttest Administration

Baseline characteristic	Control Group		Experimental Group	
	n = 34	%	*n* = 26	%
Concern				
Yes	32	94	15	58
No	2	6	11	42
Academic				
Yes	29	85	19	73
No	5	15	7	27
Climate				
Yes	27	79	17	65
No	7	21	9	35

Fisher's exact tests were conducted to compare participants in the control and experimental group on their level of professional concern about equity in school discipline, if equity in school discipline affected their students academically, and if equity in school discipline affected the climate of their school. Participants indicated their level

of concern before taking the pretest TMAS and after taking the posttest TMAS. The following items are reported for each Fisher's exact test: the number and percentage of teachers who selected yes or no, the exact *p* value, 95% CIs, and the odds ratio (*OR*) as an effect size measure for both significant and nonsignificant results. An *OR* of 1.49 or less indicates a small effect size, *OR* = 3.45 a medium effect size, and *OR* = 9 or greater a large effect size.

Before taking the pretest TMAS, a Fisher's exact test indicated the percentage of participants concerned about the topic of school discipline was not significantly different by the experimental group, 81% (21/26), and control group, 91% (31/34) (p = .43, 95% CI [0.42, 17.3], *OR* = 2.42). These results indicate a small to medium practical effect size between groups. After taking the posttest TMAS, the percentage of participants concerned about the topic of equity in school discipline was significantly different by the experimental 58% (15/26) and control group 94% (32/34) (p = .001, 95% CI [2.08, 117], *OR* = 11.2). This indicates a large effect size between the experimental and control group. Experimental group participants were significantly less concerned about the topic of equity in school discipline than the control group after taking the posttest TMAS.

Before taking the pretest TMAS, the percentage of participants who thought equity in school discipline affected their students academically was not significantly different by the experimental group, 81% (21/26), and the control group, 85% (29/34) (p = .73, 95% CI [0.28, 6.82], *OR* = 1.37). This indicates a small practical effect size between the control and experimental group. After taking the posttest TMAS, the percentage of participants who thought equity in school discipline affected their students academically was not significantly different by the experimental group, 73% (19/26), and

the control group, 85% (29/34) (p = .33, 95% CI [0.49, 9.77], OR = 2.11). This indicates a small practical effect size between the control and experimental group.

Before taking the pretest TMAS, the percentage of participants who thought equity in school discipline affected the climate of their school was not significantly different by the experimental group, 77% (20/26), and the control group, 82% (28/34) (p = .75, 95% CI [0.32, 6.07], OR = 1.39). This indicates a small practical effect size between the experimental and control group. After taking the posttest TMAS, the percentage of participants who thought equity in school discipline affected the climate of their school was not significantly different by the experimental, 65% (17/26), and the control group, 79% (27/34) (p = .25, 95% CI [0.55, 7.72], OR = 2.02). This indicates a small practical effect size between the control and experimental group.

Table 6

Pretest Teacher Multicultural Attitude Survey (TMAS) Item Analysis

	Number and Percentage of Responses by Item					*Mdn*	*M*	*SD*
	1	2	3	4	5			
1. I find teaching a culturally diverse student group rewarding.	0 (0)	0 (0)	2 (3)	22 (37)	36 (60)	5	4.6	0.56
2. Teaching methods need to be adapted to meet the needs of a culturally diverse student group.	0 (0)	1 (2)	4 (7)	24 (40)	31 (52)	5	4.4	0.70
3[a]. Sometimes I think there is too much emphasis placed on multicultural awareness and training for teachers.	6 (10)	36 (60)	8 (13)	8 (13)	2 (3)	2	3.6	0.96
4. Teachers have the responsibility to be aware of their students' cultural backgrounds.	1 (2)	1 (2)	1 (2)	29 (48)	28 (47)	4	4.4	0.76
5. I frequently invite extended family members (e.g., cousins, grandparents, godparents, etc.) to attend parent teacher conferences.	3 (5)	18 (30)	17 (28)	13 (22)	9 (15)	3	3.1	1.15
6[a]. It is not the teacher's responsibility to encourage pride in one's culture.	8 (13)	36 (60)	14 (23)	2 (3)	0 (0)	2	3.8	0.69
7. As classrooms become more culturally diverse the teacher's job becomes increasingly challenging.	0 (0)	5 (8)	9 (15)	29 (48)	17 (28)	4	4.0	0.88
8. I believe the teacher's role needs to be redefined to address the needs of students from culturally diverse backgrounds.	0 (0)	5 (8)	20 (33)	30 (50)	5 (8)	4	3.6	0.77
9. When dealing with bilingual students, some teachers may misinterpret different communication styles as behavioral problems.	0 (0)	1 (2)	4 (7)	39 (65)	16 (27)	4	4.2	0.62
10. As classrooms become more culturally diverse, the teacher's job becomes increasingly rewarding.	0 (0)	0 (0)	9 (15)	35 (58)	16 (27)	4	4.1	0.64
11. I can learn a great deal from students with culturally different backgrounds.	0 (0)	0 (0)	3 (5)	22 (37)	35 (58)	5	4.5	0.60
12[a]. Multicultural training for teachers is not necessary.	20 (33)	32 (53)	4 (7)	4 (7)	0 (0)	2	4.1	0.81
13. In order to be an effective teacher, one needs to be aware of cultural differences present in the classroom.	0 (0)	1 (2)	0 (0)	32 (53)	27 (45)	4	4.4	0.59
14. Multicultural awareness training can help me work more effectively with a diverse population.	0 (0)	1 (2)	1 (2)	38 (63)	20 (33)	4	4.3	0.58
15[a]. Students should learn to communicate in English only.	12 (20)	34 (57)	9 (15)	5 (8)	0 (0)	2	3.9	0.83
16[a]. Today's curriculum gives undue importance to multiculturalism and diversity.	4 (7)	33 (55)	16 (27)	6 (10)	1 (2)	2	3.5	0.83
17. I am aware of the diversity of cultural backgrounds in my classroom.	0 (0)	1 (2)	3 (5)	40 (67)	16 (27)	4	4.2	0.60
18. Regardless of the racial and ethnic makeup of my class, it is important for all students to be aware of multicultural diversity.	0 (0)	1 (2)	0 (0)	33 (55)	26 (43)	4	4.4	0.59
19[a]. Being multiculturally aware is not relevant for the subject I teach.	24 (40)	30 (50)	5 (8)	1 (2)	0 (0)	2	4.3	0.69
20[a]. Teaching students about cultural diversity will only create conflict in the classroom.	33 (55)	22 (37)	2 (3)	2 (3)	1 (2)	1	4.4	0.85

Note. 1 = Strongly Disagree, 2 = Disagree, 3 = Uncertain, 4 = Agree, 5 = Strongly Agree.
[a] Items are reverse scored.

Table 7

Posttest Teacher Multicultural Attitude Survey (TMAS) Item Analysis

	Number and Percentage of Responses by Item					*Mdn*	*M*	*SD*
	1	2	3	4	5			
1. I find teaching a culturally diverse student group rewarding.	0 (0)	0 (0)	2 (3)	30 (50)	28 (47)	4	4.4	0.56
2. Teaching methods need to be adapted to meet the needs of a culturally diverse student group.	0 (0)	3 (5)	3 (5)	30 (50)	24 (40)	4	4.2	0.77
3[a]. Sometimes I think there is too much emphasis placed on multicultural awareness and training for teachers.	8 (13)	31 (52)	16 (27)	5 (8)	0 (0)	2	3.7	0.81
4. Teachers have the responsibility to be aware of their students' cultural backgrounds.	1 (2)	1 (2)	2 (3)	37 (62)	19 (32)	4	4.2	0.73
5. I frequently invite extended family members (e.g., cousins, grandparents, godparents, etc.) to attend parent teacher conferences.	4 (7)	22 (37)	14 (23)	14 (23)	6 (10)	3	2.9	1.13
6[a]. It is not the teacher's responsibility to encourage pride in one's culture.	17 (28)	27 (45)	12 (20)	3 (5)	1 (2)	2	3.9	0.92
7. As classrooms become more culturally diverse the teacher's job becomes increasingly challenging.	0 (0)	6 (10)	8 (13)	36 (60)	10 (17)	4	3.8	0.83
8. I believe the teacher's role needs to be redefined to address the needs of students from culturally diverse backgrounds.	0 (0)	11 (18)	14 (23)	28 (47)	7 (12)	4	3.5	0.93
9. When dealing with bilingual students, some teachers may misinterpret different communication styles as behavioral problems.	0 (0)	0 (0)	3 (5)	40 (67)	17 (28)	4	4.2	0.53
10. As classrooms become more culturally diverse, the teacher's job becomes increasingly rewarding.	0 (0)	0 (0)	8 (13)	32 (53)	20 (33)	4	4.2	0.66
11. I can learn a great deal from students with culturally different backgrounds.	0 (0)	0 (0)	1 (2)	30 (50)	29 (48)	4	4.5	0.54
12[a]. Multicultural training for teachers is not necessary.	25 (42)	25 (42)	9 (15)	1 (2)	0 (0)	2	4.2	0.77
13. In order to be an effective teacher, one needs to be aware of cultural differences present in the classroom.	0 (0)	1 (2)	1 (2)	31 (52)	27 (45)	4	4.4	0.62
14. Multicultural awareness training can help me work more effectively with a diverse population.	0 (0)	2 (3)	3 (5)	34 (57)	21 (35)	4	4.2	0.70
15[a]. Students should learn to communicate in English only.	17 (28)	25 (42)	12 (20)	5 (8)	1 (2)	2	3.9	0.98
16[a]. Today's curriculum gives undue importance to multiculturalism and diversity.	6 (10)	30 (50)	17 (28)	4 (7)	3 (5)	2	3.5	0.95
17. I am aware of the diversity of cultural backgrounds in my classroom.	0 (0)	2 (3)	4 (7)	40 (67)	14 (23)	4	4.1	0.66
18. Regardless of the racial and ethnic makeup of my class, it is important for all students to be aware of multicultural diversity.	0 (0)	1 (2)	0 (0)	39 (65)	20 (33)	4	4.3	0.56
19[a]. Being multiculturally aware is not relevant for the subject I teach.	23 (38)	31 (52)	5 (8)	1 (2)	0 (0)	2	4.3	0.69
20[a]. Teaching students about cultural diversity will only create conflict in the classroom.	25 (42)	32 (53)	3 (5)	0 (0)	0 (0)	2	4.4	0.58

Note. 1 = Strongly Disagree, 2 = Disagree, 3 = Uncertain, 4 = Agree, 5 = Strongly Agree.
[a] Items are reverse scored.

Experimental group participants were asked to complete the race IAT at time points 1, 2, and 3. Of the 39 participants assigned, 27 participants completed the IAT at time point 1, 15 participants at time point 2, and seven participants at time point 3 (see Table 8). IAT D-scores are bound at ±2. A score of +2 would indicate a strong preference for White skin and a score of -2 a strong preference for Black skin. The descriptive statistics for participants who completed all three IAT assessments are reported in Table 10 and demographic information for the seven participants who completed all three IAT's are reported in Table 9. Descriptive statistics for participants who scored low or high on the pretest TMAS, along with their posttest IAT scores, are presented in Table 11. Only participants who completed the IAT at time point 3 are included in the quantitative results of questions 2 and 3.

Table 8

Descriptive Statistics of all IAT D-Scores at Time Points 1, 2, and 3

Time Point	*n*	*M*	*SD*	Min[a]	Max[b]	Skewness	Kurtosis
IAT1	27	0.36	0.30	-0.36	0.9	-0.61	-0.01
IAT2	15	0.18	0.49	-0.59	1.17	0.38	-0.58
IAT3	7	-0.04	0.49	-1.03	0.5	-0.96	-0.38

Note. [a]Minimum value; [b]Maximum value.

Table 9

Demographic Characteristics of Seven Participants who Completed the IAT at Time Points 1, 2, and 3

Participant	1	2	3	4	5	6	7
Age	26-34	35-44	35-44	45-54	35-44	26-34	45-54
Gender	Female	Female	Female	Male	Female	Female	Female
Race	White	White	White	NA[d]	White	White	White
Experience[a]	8-12	13-18	19+	19+	13-18	8-12	19+
Current Assignment[b]	High School	Elementary	Elementary	Elementary	Elementary	Elementary	Elementary
Highest Degree[c]	Master's	Bachelor's	Specialist's	Specialist's	Specialist's	Bachelor's	Master's
Pretest TMAS	71	91	89	73	76	74	74
Posttest TMAS	76	95	85	NA	79	77	74
Pretest Concern	No	Yes	Yes	NA	Yes	Yes	Yes
Posttest Concern	Yes	No	Yes	NA	Yes	Yes	Yes
Pretest Academic	Yes	Yes	No	NA	Yes	Yes	Yes
Posttest Academic	Yes	Yes	No	NA	Yes	Yes	No
Pretest Climate	Yes	Yes	No	Yes	Yes	Yes	Yes
Posttest Climate	Yes	Yes	No	NA	Yes	Yes	Yes
IAT[e]1	0.41	-0.36	0.39	0.33	0.45	0.79	0.90
IAT[e]2	-0.14	-0.03	1.17	0.22	-0.59	-0.51	0.48
IAT[e]3	0.24	-1.03	0.12	0.07	0.05	-0.24	0.50

Note. [a]Total years of educational experience; [b]Current teaching assignment; [c]Highest Degree Earned; [d]No Response; [e]IAT D-Score

Table 10

Descriptive Statistics of Participant Repeated Measure D-Scores on the IAT

Time	*n*	*M*	*SD*	Min[a]	Max[b]	Skewness	Kurtosis
IAT1	7	0.42	0.41	-0.36	0.9	-0.64	-0.69
IAT2	7	0.09	0.61	-0.59	1.17	0.51	-1.19
IAT3	7	-0.04	0.49	-1.03	0.5	-0.96	-0.38

Note. [a]Minimum value; [b]Maximum value.

Table 11

Descriptive Statistics of Mean Pretest TMAS Scores and Mean IAT3 D-Scores by Low and High Groups

Group	*n*	*M*	*SD*	Min	Max	Skewness	Kurtosis
Low (TMAS)	4	73	1.41	71	74	-0.53	-1.88
High (TMAS)	3	85.33	8.14	76	91	-0.36	-2.33
Low (IAT3)	4	0.14	0.31	-0.24	0.50	-0.38	-1.99
High (IAT3)	3	-0.29	0.65	-1.03	0.12	-0.38	-2.33

Note. [a]Minimum value; [b]Maximum value. Participants were grouped based on pretest TMAS scores.

Quantitative Results

Results by Question

1. Is there a significant difference between pretest and posttest scores on Ponterotto's (1995a) Teacher Multicultural Attitude Survey (TMAS) by the control group and experimental group?

Descriptive statistics indicated the overall mean TMAS pretest score for the (N = 60) participants who completed the pretest and posttest TMAS was M = 82 (SD = 7.32). The overall mean TMAS posttest score was M = 81 (SD = 8.19). Control group participants pretest TMAS scores (n = 34) ranged from 68 – 99 with a mean score of 83.88 (SD = 6.29) and posttest scores ranged from 66–97 with a mean score of 82.74 (SD = 6.93). Experimental group participants pretest TMAS scores (n = 26) ranged from 57 - 91 with a mean score of 79.54 (SD = 7.94) and posttest scores ranged from 54 - 96 with a mean score of 78.73 (SD = 9.23) (see Table 12).

Table 12

Descriptive Statistics of Control and Experimental Group Mean TMAS Scores by Pretest and Posttest

Group	Time	*n*	*M*	*SD*	Min[a]	Max[b]	Skewness	Kurtosis
Control	Pretest	34	83.88	6.29	68	99	0.01	0.07
Experimental	Pretest	26	79.54	7.94	57	91	-0.64	0.4
Control	Posttest	34	82.74	6.93	66	97	-0.21	-0.26
Experimental	Posttest	26	78.73	9.23	54	96	-0.16	0.25

Note. [a]Minimum value; [b]Maximum value.

A one-way ANCOVA was conducted to compare the mean pretest and posttest TMAS scores of the control group ($n = 34$) and experimental group ($n = 26$) after controlling for group mean pretest TMAS scores. Statistical assumptions and considerations were checked before the analysis was conducted. There was one missing data for item 4 on the pretest TMAS. The decision was made to enter the median score of all other participants at time point 1 for that participant's pretest TMAS item 4 score. Data were converted to z-scores to examine outliers. An outlier was identified with a value greater than 3 ($z = -3.16$). The decision was made to keep this score. Skewness and kurtosis values indicated the data were normally distributed (see Table 12).

It was determined there was a linear relationship between pretest and posttest TMAS scores after visual inspection of the scatterplot (see Figure 2). The assumption of normally distributed scores was tested statistically utilizing the Heteroscedasticity-Corrected Covariance Matrices (Fox, 2016). The assumption of homogeneity of regression slopes was met as the interaction was not statistically significant, $F(1, 56) = 0.05$, $p = .82$. A Shapiro-Wilk test was then conducted to check the normality assumption and was nonsignificant $W(56), = 0.99$, $p = .78$, indicating the data sets had a normal distribution. Levene's test was conducted and was also nonsignificant $F(1, 58) = 2.14$, $p = .15$, indicating the assumption of homogeneity of residual variances among the

groups were met. There were no standardized residuals with an absolute value equal to or greater than three, indicating there were no outliers in the groups. After adjusting for pretest TMAS scores, the results of the one-way ANCOVA found there was no significant difference in posttest TMAS scores between the control and experimental groups, $F(1, 57) = 0.07$, $p = .80$, $\eta p^2 = .001$. The generalized eta squared (.001) value indicated a small practical effect size.

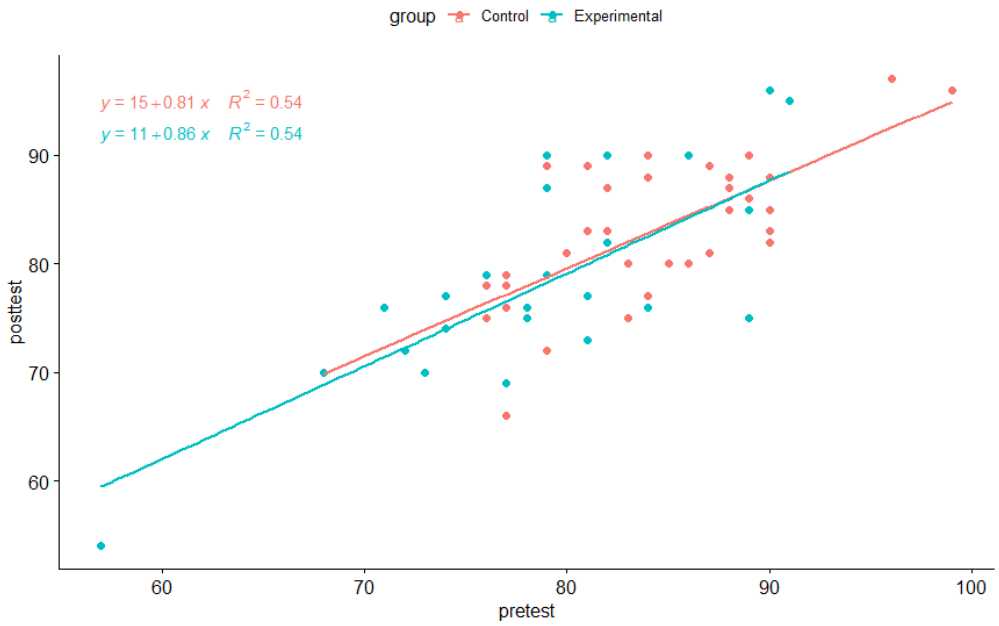


Figure 2. Scatterplot between covariate pretest scores and the outcome variable (posttest) with regression lines.

2. Is there a significant difference among participants who score low or high on pretest TMAS on the final IAT score?

A one-way analysis of variance (ANOVA) was performed to determine if participants who scored low or high on the pretest TMAS had significantly different scores on the final Implicit Association Test (IAT3). Only experimental group participants (n =7) who completed all three IAT administrations were included in the

analysis. This group of participants included six White females and one male who did not report race. Most of the teachers (six), taught at the elementary level and had over 15 years of experience in the classroom. The teachers also included two with a bachelor's degree, two with a master's degree, and three reported holding an Educational Specialist's degree.

Descriptive statistics indicated the overall mean posttest IAT score for the seven participants who completed the study was M = -0.04 (SD = 0.49). The data was ordered to separate the participant pretest TMAS scores into 2 groups between low (M = 73, SD = 1.41) and high (M = 85.33, SD = 8.14). This created the comparison groups (see Table 11). The mean IAT3 D-scores for the low group (n = 4) was M = 0.14 (SD = 0.31) and high group (n = 3) was M = -0.29 (SD = 0.65).

Before running the analysis, the data were checked to ensure statistical considerations and assumptions were met. There were no missing data. No outliers were identified. Skewness values were all within the normal range, kurtosis values were slightly platykurtic. Visual inspection of Q-Q plots indicated the scores were normally distributed. A Shapiro-Wilk test was conducted to meet the normality of residuals assumption and was nonsignificant for both groups $W(7) = 0.89$, $p = .28$, the low group $W(4) = 1.0$, $p = .99$, and the high group $W(3) = 0.80$, $p = .10$. Levene's test was conducted and the assumption of equal variances was met, $F(1, 5) = 0.25$, $p = .63$.

The results of the ANOVA indicated there was no significant difference in mean IAT3 D-scores between participants who scored low or high on the pretest TMAS ($F(1, 5) = 1.41$, $p = .29$, $\eta p^2 = .22$). The generalized eta squared value indicated a moderate practical effect size. Due to the small sample size of participants who completed the

study, the effect size is included to share practically significant results. Despite there being no statistical difference between IAT scores, further analysis using Cohen's *d* was conducted to examine the effect size between groups. The effect size ($d = 0.91$) indicated a large practical effect in mean IAT3 D-scores between participants who scored low and high on the pretest TMAS.

3. Is there a significant difference in participant repeated measure scores on the Implicit Association Test (IAT)?

A repeated-measures ANOVA was performed to determine if participant IAT scores changed over time. Before running the analysis, statistical considerations and assumptions were checked. There were no missing data. One outlier was identified but after examining the data the score was consistent for the participant. The mean for time point 1 was $M = 0.42$ ($SD = 0.4$), time point 2 $M = 0.09$ ($SD = 0.61$), and time point 3 $M = -0.04$ ($SD = 0.49$), respectively (see Table 10). Skewness values were all within the normal range. One kurtosis value (IAT2) was slightly platykurtic. A visual inspection of Q-Q plots and histogram indicated the data were normally distributed (see Figure 3). A Shapiro-Wilk test was conducted and was nonsignificant for time point 1, $W(7) = 0.89$, $p = .26$, time point 2, $W(7) = 0.94$, $p = .65$, and time point 3 $W(7) = 0.86$, $p = .14$ to completely meet the normality assumption. Mauchly's test was conducted and was nonsignificant ($W = .42$, $p = .11$), indicating the variances of group differences were equal meeting the sphericity assumption. The results of the repeated measures ANOVA indicated there was no statistically significant difference in mean repeated measure IAT scores, $F(2, 12) = 2.23$, $p = .15$ (see Figure 4). Due to a limited number of participants, however, further analysis using Cohen's *d* was conducted to analyze the effect size of

time on IAT scores to determine any practically significant results. A large practical effect size (d = 1.54) between time points 1 and 3 was observed. A small to medium practical effect size was observed (d = 0.45) when comparing the scores from time point 1 and 2 and a small practical effect size (d = 0.2) between time points 2 and 3.

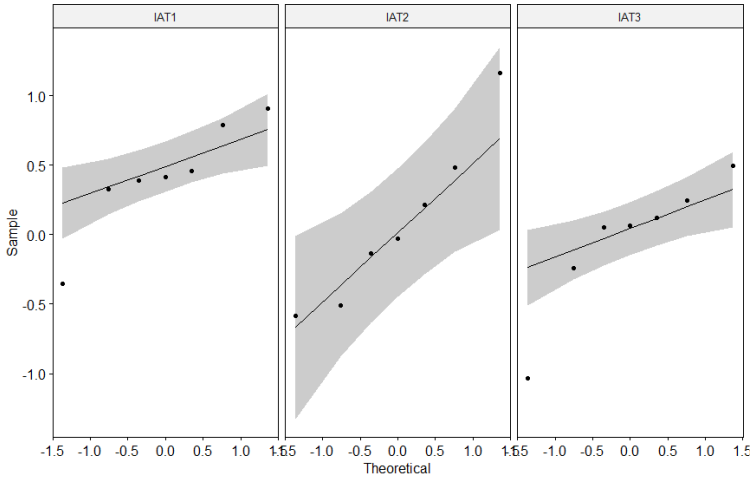


Figure 3. Q-Q plots of residuals by IAT d scores at time points 1, 2, and 3, respectively.

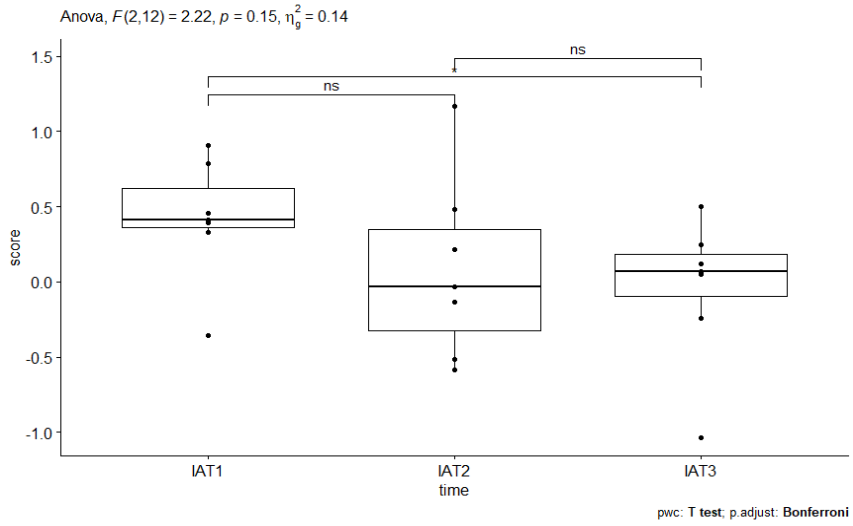


Figure = 4. Box plots with p value for the results of the repeated measures ANOVA

Qualitative Results

The qualitative component of this study was designed to better understand and explain the quantitative results generated by the licensed teachers in this study. Creswell and Plano Clark (2017) suggest using the explanatory sequential design help explain quantitative results by allowing participants to describe their experience as a participant. Onwuegbuzie and Collins (2007) contend sharing quantitative results with the participants provide opportunities to gather both confirming and/or disconfirming evidence of the quantitative findings. Participants from both the control and experimental group were included in this portion of the study. Gathering data from both groups and comparing their different experiences were critically important to better understand how to broach this topic with licensed teachers moving forward. Creswell and Plano Clark explain crafting qualitative questions based on the analysis and results of the quantitative phase is a sound strategy for minimizing threats to the validity of the study. Questions were drafted based on the quantitative results. Two expert reviewers analyzed the quantitative results and questions. The feedback provided by the expert reviewers helped clarify and simplify interview questions (see Appendix H). The questions were analyzed once more by a panel of expert reviewers as a validation measure (see Appendix I). Based on panel feedback, questions were edited to remove bias, ambiguity, and technical jargon. Leading questions were eliminated or revised. One question regarding K-12 leadership was added so participants could clearly articulate how this topic might best be addressed moving forward.

Participants

Participants were invited to participate at time point 3 via Qualtrics software. Volunteer participants from both the control and experimental group were selected to participate and invited via email. Five participants from each group were scheduled for an interview. One participant from the control group was unable to participate. A total of nine interviews were conducted. The first two interviewees received the revised questions beforehand. All interviews were conducted and recorded virtually via Microsoft Teams software. Social distancing guidelines amid the COVID pandemic made in-person interviews untenable. A virtual platform allowed participants to be interviewed in a setting and time comfortable for them. The first two interviewees were contacted to validate and review the presentation and their responses as a qualitative validity strategy. Creswell and Plano Clark (2017) state "member-checking is a frequently used strategy in which the investigator takes summaries of the findings back to key participants in the study and asks them if it is an accurate reflection of their experiences" (p. 217). A definition of key terms and timeline of the study was added, and the quantitative presentation was simplified so those unfamiliar with statistical analysis and interpretation might better understand the results (see Appendix E). Due to the amount of time that passed from the beginning of the study to the qualitative phase, participant feedback indicated this additional information was vital to refresh their memory on the process of participating in the quantitative phase.

Demographic Characteristics

The remaining seven interviews took place during the start of the 2020-2021 school year. Interviews were conducted after school, on weekends, and during their planning time. By utilizing a virtual approach for interviews, respondents could participate at a time and location convenient for them. The interviewees were offered the revised questions 24 hours before their scheduled interview appointment. Due to the socially sensitive nature of the topic and the underlying general unease regarding close physical proximity during a global pandemic, this approach allowed respondents to respond safely, openly, and at length.

The five participants from the experimental group were all female and White (see Table 13). All five participants held graduate degrees and had over 10 years of experience in the classroom. All five participants currently teach at the elementary level. The four participants from the control group were all White as well (see Table 14). Participants from this group included two males and two females. All four participants held graduate degrees. Two of the participants taught at an elementary school, one at a junior high, and one at the high school level. The teachers in this group ranged from 3–17 years' experience teaching at a K-12 school.

Table 13

Descriptive Statistics of Participants Interviewed from the Experimental Group

	Interview 1	Interview 2	Interview 3	Interview 4	Interview 5
Pseudonym	Charlotte	Ava	Amelia	Olivia	Aurora
Gender	Female	Female	Female	Female	Female
Race	White	White	White	White	White
Experience[a]	12 years	16 years	20 years	10 years	15 years
Setting	Elementary	Elementary	Elementary	Elementary	Elementary
Degree[b]	Master's	Master's	Master's	Master's	Specialist's

Note. [a]Total years of educational experience; [b]Highest degree earned

Table 14

Descriptive Statistics of Participants Interviewed from the Control Group

	Interview 1	Interview 2	Interview 3	Interview 4
Pseudonym	James	Mary	Jennifer	John
Gender	Male	Female	Female	Male
Race	White	White	White	White
Experience[a]	5 years	17 years	7 years	3 years
Setting	Junior High	Elementary	Elementary	High School
Degree[b]	Specialist's	Specialist's	Master's	Master's

Note. [a]Total years of educational experience; [b]Highest degree earned

Results by Question

The interviews were organized and analyzed by (experimental and control) group. This process helped identify, compare, and clarify any delineation between the groups. Maxwell (2013) contends comparing qualitative data by group (control and experimental) is an effective strategy at mitigating threats to validity–especially when participants are drawn from a homogenous population (such as classroom teachers). The interview data will be referenced and compared to the quantitative data to answer question 4. The average length of interviews was 35 minutes. Teachers were interviewed during their planning time, after school, or on the weekends during the school year so time was limited. Interviewees were offered to schedule a follow-up interview to offer any additional insight or clarification (none did).

Interview data will be presented in the order the questions were delivered and are based on recurring themes evident in the literature on implicit racial bias and equity in school discipline. The themes were: acknowledgment and awareness of bias, IAT reactivity, efforts to reduce bias, teacher desire to improve professional practice, teacher decision making in the classroom, and the likely impact implicit racial bias has in K-12 schools.

4. In what ways do the interview data of licensed teachers about their views on the importance of awareness of implicit bias in classrooms provide an explanation for any quantitative results from the IAT or TMAS?

Interview Questions

1. What stands out most in your mind when you think about participating in the quantitative portion of the study?

All five participants in the experimental group indicated taking the IAT assessment stood out most to them when reflecting on the experience. Ava stated the IAT "surprised me. I don't know if it's right or wrong, but it surprised me. I missed more than I thought." Amelia echoed her sentiment by explaining "I was surprised by how slow I was on the IAT and that I didn't get 100% of them right." Charlotte reflected on her first time taking the IAT, specifically the instrument design feature that switches what key to strike (right or left) for White or Black skin tone–"I thought I had control of it but, when it switched up, it kind of changed it a little bit in my mind I guess–I wasn't expecting it." Aurora agreed, "The test itself (stood out to her). It was a little confusing, I mean, some keys I'd accidentally hit the wrong button sometimes. You really have to think." Finally, Olivia summed it up with her response: "I would say the IAT. Once I associated the keys, because I was doing it on my phone, words with one hand and words with the other hand, it took me a minute to switch the words or the pictures."

All four of the control group participants, on the other hand, indicated their change in awareness surrounding the topic of implicit bias stood out most to them. Jennifer stated, "honestly, that there were biases I didn't even realize I had and I feel I'm pretty equitable across the board as a teacher… that was the most eye-opening part of it

to me." Mary stated, "one of my biggest fears is that I would have some hidden racism or implicit bias and not realize it." James and John both agreed the experience increased their awareness of the topic and both indicated they were still reflecting on it. John stated, "I think some of the questions made me think, I'm just trying to be aware if there is any bias (in my decisions)." Two of the four participants, Mary and John, also mentioned the IAT. Although control group participants did not take the assessment, they were given a brief description of the instrument during time point 1 of the study. Mary stated she "was very curious to see what I would have scored" and John asserted "I was kind of curious to how I would have scored (on the IAT)."

2. What are your thoughts on why the mean TMAS scores were different between the control and experimental groups? Do the results surprise you?

Three of the five experimental group participants indicated the results did surprise them. They figured experimental group TMAS scores would rise significantly after taking the IAT. Amelia stated she thought experimental group scores "would have gone up and made the overall score higher because, with me, I wanted my score to go up. I felt bad about myself thinking I might have some biases." Olivia and Aurora stated that although they were surprised by the results, they figure it had to do something with the IAT. When asked if she ever felt defensive about her IAT results Olivia stated "No, I felt angry at them (the computer keys), the buttons I had to push when the words and pictures switched." Aurora inferred control group participants were by in large assessing themselves on their implicit racial bias and she thought "people are going to know what to say." Charlotte and Ava, on the other hand, were not surprised by the discrepancy. They both figured the IAT played a role in group differences. Ava theorized many

participants were unfamiliar with the topic of implicit bias and the IAT, and people were "going to change some during the quizzes (IAT)." Charlotte thought people "can control what they are thinking" and "they know what to say" to appear unbiased. She thought the IAT may have exposed some well-hidden traits–"sometimes people can live double lives."

Three of the four participants in the control group were not surprised by the difference in group scores. They figured the IAT was the main reason group scores were different. Jennifer thought members of the control group weren't as "honest" and "maybe answered how they felt they should answer and not necessarily as honest as what they should have been." Mary thought it was the combination of taking the IAT and the civil unrest in the United States caused by the murder of George Floyd in May of 2020. She stated "A lot was going on in our country during that time. You can't help but wonder if all that came into play." James was intrigued by the differences he saw in the scatterplot of TMAS scores. He thought the control group scores were tightly concentrated around "80-90, whereas the experimental group scores range from the mid-upper 90's down below 60." He wasn't able to identify a likely cause for this anomaly between two highly homogenous groups (teachers) on an instrument specifically designed for them (TMAS).

3. Does that dropout rate difference between the control and experimental group seem reasonable/explainable to you?

Three of the five experimental group participants (Ava, Amelia, and Olivia) did not consider the dropout rate difference between the groups reasonable. Two of the three, however (Amelia and Oliva), did think the difference was explainable. Amelia

stated “I didn’t see a reason to drop out, I mean it did frustrate me (the IAT). . . I mean, that just seems like a big dropout.” She thought some participants may have gotten fearful or defensive because of the IAT. Olivia was direct with her thoughts stating “I took the IAT 3 times and I didn’t drop out. . . I guess people didn’t like getting called out (by the IAT).” Charlotte and Aurora considered the dropout difference reasonable and explainable. Aurora thought the extra time it took experimental group participants was a valid reason stating, “people probably got tired of doing it, it took more time and people don’t typically like to complete surveys.” Charlotte thought technology was largely to blame. She stated many in her community had multiple issues with broadband service in the spring of 2020.

Control group participants Mary and John considered the dropout difference reasonable and explainable. Mary stated she had investigated the IAT after reading the description during time point 1 and thought the IAT seemed “kind of personal” and considered the validity of the instrument questionable. She once again referenced the social unrest going on in Minneapolis while the study was being conducted and thought the IAT may have been too much for some participants to handle. She theorized “people were maybe scared to find out more about themselves, I don’t know.” John also considered the difference reasonable due to the IAT and maybe, “if people weren’t happy with their (IAT) score. . . they may not want to do it again.” James and Jennifer both considered the dropout rate difference unreasonable and unexplainable. James first questioned the demographics of the groups–were they similar. After hearing the groups were similar in every demographic variable, he was incredulous, stating “That’s what they (teachers) do for a living! I wonder what did cause the change.” He then made sure

he understood what the TMAS was intended to measure (teacher awareness and sensitivity to multicultural issues in the classroom). After some thought, he said "I wonder if, throughout the process (taking the IAT), they just became more culturally aware (of themselves). If they got paranoid about the answers they were picking or if it was deliberate (dropping out)."

4. What are your thoughts on this change in preference (on IAT scores)?

All five experimental group participants thought the downward trend of IAT scores was due to increased awareness and deliberate effort. Charlotte considered the multiple measures an important component of the study because "the more it's brought up, the more aware people are of it (implicit bias)." Ava considered it a teacher's "job" to "learn other people's cultures." She added that she would like to "take it (IAT) again." All the participants considered lowering their IAT score professionally important to them. Amelia offered "many (teachers) today, I mean everybody, they're trying (to become culturally aware)." Although she considered the dropout rate difference between groups concerning, she pointedly admitted "I was trying to drop my (IAT) score." Olivia and Aurora also thought the repeated measure design of the study was a strength due to the deliberate thought process involved when one is trying to lower their IAT scores.

Two of the four control group participants considered the change in IAT scores a result of effort by participants to lower their scores. James had taken an IAT after the completion of the survey and thought it would be difficult to lower his score. He stated one would have to "intentionally focus" and pay attention to the keys you are striking rather than skin tone. Jennifer spoke about her school and colleagues. She explained her school had a majority of African American and Hispanic students. The teachers at her

school "know what they are supposed to say and what they are supposed to feel." She believed if a teacher received IAT results indicating they harbored implicit racial biases, and the teacher acknowledges it and becomes aware of their bias, it "could definitely push you to change." Mary and John were less forthcoming with their thoughts on this preference change. John did not offer an opinion or theory on the change, and Mary considered the results "interesting." She referenced her post-survey research on the IAT and remains skeptical about the instrument.

5. A number of participants provided feedback this experience was *personally* significant. How did this study alter or shape your feelings on the topic of equity in school discipline? Do you think the topic is more or less important for k-12 students and teachers after participating?

Four of the five experimental group participants thought the experience was personally significant and after participating considered the topic of equity in school discipline more important. All four of these participants indicated the experience made them more aware of implicit bias. Ava, Amelia, and Olivia thought the issue of implicit bias and equity in school discipline should be addressed more often in their school. They also stated participating in the study helped them "stop and think" when they make decisions in their classroom and school. Amelia stated it bothered her when she realized she might prefer one skin tone over another, but in the end, offered that "we all need to be cognizant of it." Charlotte, on the other hand, explained she had recently completed a yearlong course on the topic of implicit racial bias that had already "opened her eyes."

All four control group participants considered the experience personally significant and made them more aware of implicit bias and equity in school discipline.

James spoke at length about some of the underlying issues he thought were by in large responsible for inequity in school discipline. Although he considered implicit racial bias the main reason for discipline disproportionality, he also thought lower parental involvement and differential treatment of African American students were factors to consider. After some consideration, he stated participating in the study reinforced his belief that if positive teacher/student relationships were prioritized in schools we could "overcome inequity." Mary considered the experience "very beneficial" and made her more aware of potential bias in her decision making. She considered her career as a "calling" and is constantly looking for ways to improve her efficacy and practice. Jennifer and John agreed, they both stated the experience made them more aware of the topic. Jennifer stated although becoming aware was "uncomfortable", she felt by acknowledging it she was able to do something about it. She stated, "I mean, I want to improve and do whatever I can to help my students." She thinks implicit racial bias plays a large part in discipline disproportionality and now thinks "it is more widespread than we originally thought." She thinks there are probably a lot of teachers across the nation in the "same boat I'm in." John stated the experience was personally significant because it made him more aware of implicit bias and how it affects his students. He also considered the experience beneficial professionally because he now better understands what his students "need" from him to develop positive teacher/student relationships.

6. How do the issues of race, culture, and bias impact your students (Blitz et al., 2016)?

Interestingly, four of the five experimental group participants thought educating students on the issues of race, culture, and bias would be beneficial. Ava thought it would help students be more tolerant and kind to their peers. Amelia thought many racial

attitudes are "taught" at home and in social settings outside the school. She also thought many schools reinforced racial stereotypes by grouping students by color. She thought this process made it easier for one group to be treated differently than the other and eventually, many teachers "look at the kids in the Black group different than the kids in the White group." Olivia thought it was important to learn about her students' culture so she could better understand what might be acceptable in their culture and what "is not acceptable at school." This background knowledge allows her to teach her students about these differences before they become an issue at school. Aurora thought that lower teacher expectations negatively affect student performance. Charlotte, however, considered the issues of race and culture overemphasized in modern society. She referenced the recent riots in Minneapolis and the Black Lives Matter campaign and thought the never-ending media coverage was exacerbating racial tension in our nation. She stated, "they just keep wanting to dig and dig and dig, and I'll be honest, I feel kind of put in the middle." She went on to explain her husband works in law enforcement and she was concerned for his safety. She stated he was a White man working as a deputy in a predominantly African American community.

All four control group participants thought the issues of race, culture, and bias had both a direct and negative impact on their students. All four of the participants also thought educating students about the topics of race, culture, and implicit bias would benefit many of the students they teach. James referenced the gang culture predominant in one of the neighborhoods near his school. He stated some of his African American students admitted to him they did not care for White people. When asked why, a few of them stated because of what their parents told them. James believes this creates a cycle

of animosity that is nearly impossible to break. Mary agreed with James stating once a teacher/student relationship begins to sour, it can affect students academically and socially for the rest of their academic career. She thinks many of her students don't necessarily understand the events that occurred over the summer (George Floyd, etc.). She stated, students "don't understand where all the rage and animosity comes from." John also mentioned the impact of the George Floyd story. He explained he taught at the high school level and many of his students had seen the shootings over the summer. He stated all the cases on the news were those of White police officers shooting African American men and "well, they look around their school and all they see are a bunch of White teachers." Jennifer also considered the issues of race, culture, and bias affect the students at her school. She said one negative effect can create a host of other issues. She explained "lower expectations lead to increased behavior, more office discipline referrals, less patience with them, when they want or need. . . more attention? And when they don't get that attention and slip through the cracks, they won't get the chance to be successful."

7. Are there any stories or personal experiences with implicit bias in the k-12 setting you can recall? Are you willing to describe an experience or event at a school you worked at related to implicit bias and equity in school discipline? How did this experience affect you? How did it affect the student?

Experimental group participants Ava, Amelia, and Olivia were able to recall personal experiences related to implicit bias and equity in school discipline. Charlotte and Aurora were unable to think of any during the interview. Ava, Amelia, and Olivia stated their stories affected them deeply and influence their classroom decisions on a

near-daily basis. Ava told the story of an African American male she had in her class several years ago. He was often off task and did not get along with his peers. She remembers how her patience with him was getting shorter. She found herself getting onto him more often and for minor disruptions. She made a conscious effort to slow down and take time to listen to him. Over the next few weeks, the student's behavior completely changed (for the better). Amelia's story centered on her student teaching experience long ago. She cannot remember if it was her professor or mentor teacher that pointed it out to her–but they helped her see a bias she had against one of her students. And once she realized and acknowledged it, it affected her deeply. She said "I've been thinking about it for years. You know? I mean it's never left me." Olivia offered a similar experience when she said something to correct a child in her class years ago. She stated later in her career she realized she was in the wrong, that the situation was "different than what I thought initially." Although she did not share specific details, she did say once she became aware of it, she wished she "could go back and change that or say something different or do something different. I mean, I wonder how did it end up affecting that child? Did it affect that student's outcome?"

All four control group participants were able to recall a personal experience related to implicit bias and equity in school discipline. James recalled two similar instances and explained how different approaches to school discipline led to different student outcomes. Both instances were about a fight between students. In the first story, the principal threatened the student with a paddle if he misbehaved again. In two weeks, the student was back in the office and received corporal punishment. The student's behavior became worse after the paddling and he spent a great deal of time out of the

classroom for discipline the remainder of the school year. In the second instance, at a different school, the principal called the students in his office and listened to the students. When the students told their story, the principal said he trusted them and sent them back to class. After reviewing the video of the fight, the principal realized one of the students had lied. He called him back into the office and told the student he "cheated on him" . . . and they both started laughing (principal and student). The student admitted his transgression and accepted the punishment. For the rest of that school year, the student and principal spoke regularly, and a relationship was formed. He was no longer a behavior problem at school. Mary recalled a time she had to write a student up for having marijuana at school. She did not want to write him up but was forced to because of school policy. The student was sent to in-school suspension (ISS) and immediately started struggling. Although she did not know the student well, she started calling ISS during her class and asking for him to come to her class. Once the student realized she genuinely cared for him, a relationship developed that improved the teaching and learning in her class. Jennifer remembered her first-year teaching. She got a teaching position at a school with a student population that was 95% African American. She was completely unfamiliar with African American culture. About two weeks into the school year (she was teaching Kindergarten), one of her students blurted out a cuss word. She remembered saying "that's bad! we don't say those words, that's bad." After a moment, the student replied "well, my mom is bad then." She went on to say, "I really learned that language and culture have to be approached in a way that is school appropriate and respectful of their family dynamics." John shared a story about an instance near the end of a class. He had taken a student's cell phone and the student was demanding it back.

When he wouldn't give it to him, a shouting match erupted, and the student stormed out of class and bumped against him. After reflecting, he realized his part in the escalation of events–and now wonders if implicit bias played a role in the situation. He thinks many situations or instances at the high school that escalate quickly are due to decisions influenced by implicit bias.

8. How should K-12 leaders address the issue of equity in school discipline with their teachers?

Charlotte and Ava were not asked this question. They were the first two interviewees. They mentioned the role school leadership had in reducing or eliminating discipline disproportionality. Their feedback helped validate the qualitative questions and they recommended adding this question. The other three experimental group participants all thought school leaders should address the issue of equity in school discipline by developing honest relationships and having meaningful conversations with their teachers. Amelia thought leaders could also recommend their teachers take something like the race IAT. She stated she recently took the politics IAT and thinks using an IAT other than race would be less threatening to teachers. Olivia stated, just like teachers need to develop relationships with their students, principals need to understand where their teachers are coming from to understand their cultural beliefs.

All four control group participants thought leaders need to develop relationships with their staff and have meaningful conversations about implicit bias with them. James thought school leaders should "pound the concept." He also thought school leaders should simply "put their foot down" on some of the issues he has noticed. He went on to say once school leaders acknowledge and are aware of implicit racial bias in their school,

they have to prioritize the issue and correct it. Mary thought leaders should have meaningful conversations and put policies and procedures in place at the school level designed to mitigate the effects of implicit bias in teacher discipline decisions. Jennifer thinks the first step is making sure school leaders are aware of implicit bias and the role it plays in their school. She also thinks it would help if leaders stopped and thought about what things look like to students, "like a day in their life." John thinks leaders should monitor student discipline data and if they notice a teacher is having "issues" with race and discipline, they should go to them directly and figure out what is going on. He thinks many discipline issues in classrooms are often deeply personal to the teacher and should be "addressed individually–not those blanket emails or blanket statements in faculty meetings." He stated, "it's almost like students in the class and relationships (teachers and leaders). Taking the time to foster and develop them, otherwise, it falls on deaf ears."

9. Would training in this area affect the teaching and learning at your school? Would training in this area affect the teaching and learning in your classroom?

Although all five experimental group participants agreed it would be beneficial, Amelia and Olivia thought it would have to be voluntary. Charlotte thought the training should center more on where our students are coming from rather than how teachers should change. Ava stated the training was needed badly, but teachers "are always so busy" and are not given adequate time to reflect on professional development activities already in place. Ava thought participating in this study was professionally beneficial and would help her make decisions in her classroom on a "case by case basis." Amelia agreed, she thinks teachers have too much mandatory training. She thought if training in

this area were voluntary it would be extremely beneficial to teachers who want to improve or change. On the other hand, she stated "there's [*sic*] people (teachers) who need it who don't know they have these biases but don't want to change." Olivia agreed with Amelia that training in this area would be beneficial at her school, but it would have to be carefully delivered. She did not know if making the training mandatory or voluntary would make a difference, stating "I don't know if there's a wrong way or a right way (to train teachers), but there definitely needs to be a way." Both Olivia and Aurora believe it all goes back to the relationships in the building between school leaders and teachers. To foster meaningful change and help all staff in the building become aware of implicit bias, relationships need to be honest and built on trust.

All four control group participants agreed training in this area is needed at their school. James believes the topic of implicit bias and equity in school discipline is the most important issue in K-12 education today. He thinks school leaders first have to acknowledge the issue and pay close attention to the impact it has on the students in their school. James went on to say he thinks it would be relatively simple to correlate discipline disproportionality with "lower reading, lower Lexile levels, lower achievement, higher poverty, higher teen pregnancy, the list goes on." Mary believes training in the area would be beneficial to any school working to improve. Jennifer indicated the teaching staff at her school were currently taking professional development courses on the topic. The teachers and leaders at her school have spent time building relationships and she found the professional learning "hopeful." None of Jennifer's colleagues appeared to get defensive and seemed to genuinely seek an understanding of how implicit bias may be affecting the teaching and learning in their classroom and

school. John thinks training in this area would benefit both new and veteran teachers alike, basically any teacher seeking ways to be more effective in their classroom.

The nonsignificant results of questions 1, 2, and 3 were presented to all interviewees to better understand and explain research question 4. Every participant in both the control and experimental group agreed that taking the IAT (or not) was powerful and likely responsible for the differential dropout rate difference between groups. Several interviewees were not familiar with analyzing and interpreting quantitative statistical results. Careful attention was given to help them understand the analysis of TMAS pretest and posttest scores by group, IAT score differences by low and high TMAS pretest scores, and IAT repeated measure scores. Interview question 1 was intended to capture participants' most poignant thoughts on their experience. All the experimental group participants mentioned taking the IAT stood out most to them. Interestingly, two of the participants in the control group mentioned reading a brief description of the IAT at time point 1 was something they remembered clearly. Both participants spent time studying the IAT so they could better understand how it works and what it is intended to measure.

Interview questions 2, 3, and 4 were designed to gather teacher perception of group differences in TMAS score by group and the drop of IAT scores over time. The majority of participants in the experimental group were surprised TMAS scores from their group did not rise significantly after taking the IAT. Two experimental group participants were quick to point out how receiving their IAT score and feedback could have caused participants to get defensive. Three of the four control group participants also figured taking the IAT was the cause of any group difference in TMAS scores.

Participants from both groups thought some teachers may have answered how they thought they should answer rather than answer honestly (on the TMAS). They figured IAT results may have caused more experimental group participants to answer honestly than control group participants due to the direct nature of IAT feedback.

One of the more interesting results from this study was the differential dropout rate between the control and experimental group. Although nearly all participants agreed taking the IAT was the likely cause of the difference, participant feedback on if this difference seemed reasonable or explainable was all over the place. Five participants thought the difference was not reasonable. Three of those five participants, however, did think the dropout was explainable. They figured participants had some type of adverse reaction to their IAT feedback that caused them to either stop taking the IAT or drop out of the study. Four participants thought the dropout rate was reasonable and explainable. Other than defensive reactions, they thought maybe participants had either technical difficulties with the IAT or simply tired of taking the assessment multiple times.

When participants were shown how IAT scores dropped over time nearly every interviewee agreed this drop in IAT scores was likely due to intentional effort (to drop their scores). Three experimental group participants considered it part of their job as a teacher to lower their IAT score and two control group participants indicated they thought that is why the scores dropped (the participants wanted them to drop). While one control group participant questioned the validity of the IAT considering how the scores did drop over time, she did think the results were "interesting." One experimental group participant indicated the IAT did make you "sit there and think", she was curious if different IAT's could be used (other than race) in future iterations of this study.

The overwhelming response from interview participants indicated by participating in this study their overall awareness of implicit bias and the importance of the topic of equity in school discipline was enhanced. Many participants thought it was an important topic to discuss and study; they found the information they garnered by participating professionally beneficial. Many participants agreed with the theory put forth by Plant and Devine (2009): Individuals must first be made aware of their bias, acknowledge how this bias may affect their decisions, and deliberately work to reduce the impact of bias if they want to change.

Summary

The results reported in this chapter followed the characteristics and order inherent of a sequential explanatory design. Quantitative results were first reported and followed by the qualitative results. The major findings of this study indicate teacher's awareness of and sensitivity to multicultural issues in the classroom as measured by the TMAS are unlikely to change in pretest and posttest measures. All participants included in the quantitative results of this study were introduced to an intervention specifically designed to increase an individual's awareness of implicit bias and were given strategies to reduce or eliminate the impact this bias may have on their professional practice. After adjusting for pretest score differences, there was no significant difference between participant TMAS scores by the control or experimental group. Experimental group participants took the race IAT. There was no significant difference in participants who scored low or high on the pretest TMAS on IAT scores taken at time point 3. Lastly, although IAT scores did trend down over time, there was no significant difference in repeated measure IAT scores.

The major findings of the qualitative portion of this study indicated although the quantitative results of this study were nonsignificant, many participants did find the experience personally and professionally significant. They all reported increased levels of awareness concerning implicit bias and think this change will make them more effective teachers. Nearly every participant thought the Race IAT was the most powerful and controversial part of the study and substantiate Howell et al.'s (2014) assertion researchers should use the Race IAT with caution. Even though all participants indicated they were concerned about equity in school discipline and implicit bias when they were recruited to participate, many interviewees thought the direct nature of IAT feedback coupled with the racial tension present during the summer of 2020 may have caused more experimental group participants to either stop taking the IAT or drop out of the study completely. The major findings from the control and experimental group interview participants were the fact teachers who want to improve their practice will go to extraordinary lengths to meet this goal.

Chapter V

DISCUSSION

This study sought to examine the topic of equity in school discipline from the perspective of classroom teachers. The purpose of this study was binary in nature. The first was to investigate the impact Devine et al.'s (2012) intervention and the Race Implicit Association Test (IAT) had on volunteer in-service teachers concerned about the topic of equity in school discipline. Teachers completed Ponterotto's (1995a) Teacher Multicultural Awareness Survey (TMAS) as a pretest and posttest measure. The second purpose was to examine participant feedback and perception of the study. Stated differently, did teachers find participating in the study professionally beneficial and did the process alter their views on the topic of equity in school discipline.

Literature Review

Equity in school discipline (discipline disproportionality) is a problem schools in the United States have grappled with for decades (Children's Defense Fund, 1975; Gordon et al., 2000; U.S. Department of Education, 2016a). Discipline disproportionality is the systemic overrepresentation of a particular subgroup or ethnicity of students when analyzing school, district, or state level discipline data (U.S. Department of Justice & U.S. Department of Education, 2014). Although discipline disproportionality can be measured in several ways, one fact remains–African American students are three to four times more likely to receive an office discipline referral (ODR) than their European American peers (Brown & Steele, 2015; Freeman & Steidl, 2016; Krezmien et al., 2006;

McIntosh et al., 2017; Okonofua & Eberhardt, 2015; U.S. Department of Education, 2016b). Students receiving an ODR inherently lose time in the classroom. This loss of instructional time has a negative effect on student achievement and increases the risk a student will drop out of high school (Bleyaert, 2009; Fisher, 2011; Gass & Laughter, 2015; Gordon et al., 2000; Irby, 2013; Monahan et. al., 2014; Nance, 2016; Noltemeyer et al., 2015; Shollenberger, 2015). More to the point, punitive discipline practice has been linked to higher rates of involvement with the juvenile justice system and increases the likelihood of being incarcerated as an adult (Nance, 2016; Nicholson-Crotty et al., 2009; Noltemeyer et al., 2015; Shollenberger, 2015). Nicholson-Crotty et al. (2009) found African American students in Missouri were more likely to face punitive discipline measures than their European American peers, even after controlling for environmental differences like poverty and urban density. Skiba et al. (2000) found after controlling for socioeconomic status and gender, African American males were more likely than their European American peers to receive an ODR–often for subjective reasons. Skiba et al. theorized this discrepancy may be due to systemic cultural biases occurring at the classroom level. Although this problem spans generations and has been studied extensively, pragmatic effective solutions remain elusive (Losen et al., 2014).

Beck and Muschkin (2012) compared behavior and achievement data for African American and European American students in North Carolina. They found as the discipline gap increased (e.g. more ODR's for Black students), so did the achievement gap. Beck and Muschkin coined the term cradle-to-prison pipeline and argued teacher perception of African American students, even at an early age, creates enduring consequences for affected students. More recently, Pearman II, Curran, Fisher, and

Gardella (2019) utilized data from the Stanford Education Data Archive and the Civil Rights Data collection. Analogous to Beck and Muschkin, Pearman II et al. (2019) found the discipline gap and the achievement gap were highly correlated for African American students. As the discipline gap between African American and European American students increased, so did the achievement gap. Additionally, Pearman II et al. found the opposite was true. African American student achievement increased as disproportionate discipline practices decreased. The results of this study support Shollenberger's (2015) research. She found the achievement gap decreased significantly for African American students after controlling for suspension. Consequently, Shollenberger found any student (African American or European American) suspended over 10 days was more likely to be arrested or drop out of high school. Pearman II et al. (2019) suggest districts working to improve the discipline gap may get the unintended benefit of closing the achievement gap (or vice versa). Fiester and Gibson's (2015) findings support this premise. They found exclusionary discipline practice was a better predictor of academic achievement than race alone.

Similarly, Sullivan and Bal (2013) found African American students were more likely to be under identified for gifted services and over identified for special education services. The U.S. Department of Education (2016a) defines special education disproportionality in the same way they define discipline disproportionality. The U.S. Department of Education utilizes a risk ratio analysis comparing the rate one particular subgroup of students has of being identified for special education services (or receiving an ODR) compared to a comparison subgroup. To that end, Sullivan and Bal believe the two are highly correlated. They assert African American students receiving exclusionary

discipline as a punitive measure inherently miss instruction, fall behind their peers academically, and subsequently endure lowered teacher expectations. Sullivan and Bal argue this cycle leads to more African American students being identified for special education services and a litany of other negative consequences (negative teacher perception, higher dropout rate, old for grade, etc.). Furthermore, Vanderhaar et al. (2014) discovered African American students entrenched in this cycle are more likely to be referred to an alternative placement or school for disciplinary reasons.

In like fashion, Hughes and Kwok (2007) found African American students were more likely to have negative teacher relationships. They suggest negative teacher/student/parent relationships were largely responsible for the achievement gap between African American students and their peers. Indeed, Mortenson (2018) reported lower teacher expectations and negative teacher perceptions of African American students were highly correlated with the achievement gap.

Reducing discipline disproportionality and eliminating the achievement gap have both been prioritized for years with limited success (Losen et al., 2014). Blitz et al. (2016) found teachers in their study were highly sympathetic to the impoverished lives many of the African American students in their classrooms face. Interestingly, the same teachers considered African American students more apathetic, aggressive, and disruptive than other students. When teachers were required to participate in cultural awareness professional development, many were insulted by the insinuation they were culturally deficient in some way (Blitz et al., 2016). Blitz et al. reported many of the teachers claimed to be colorblind while teaching–they did not consider race when reviewing student achievement and classroom discipline data. Although Sue et al. (2007) asserted

colorblindness is considered a racial micro invalidation (racist ideology), Hartmann et al. (2017) claim the colorblind ideology has evolved for many. Hartmann et al. believe individuals struggling with racial inequity and who claim to be colorblind are emotionally sensitive and should be treated with care when facing personal issues like race and culture. Consequently, forcing teachers to undergo cultural awareness training is unlikely to be an effective measure at reducing the discipline gap (Blitz et al., 2016).

Boneshefski and Runge (2014) think having teachers examine discipline data is a promising strategy to reduce discipline disproportionality. They contend by examining school level discipline data, teachers might become aware of their discipline tendencies and self-correct. Amin (2017) theorized a similar outcome for judges who examine the racial data of their courtroom decisions. To this end, helping individuals understand how implicit bias affects decision making, in a non-threatening manner, is a promising approach to reduce inequity in school discipline.

The psychological premise of implicit bias has been studied for decades (Greenwald & Banaji, 1995). Greenwald and Banaji (1995) believe when individuals are faced with making an automatic response to a given situation, implicit bias may cause them to make a different decision than they would if they had the time to carefully consider the situation. Further, Greenwald and Banaji assert many individuals are not aware of their implicit biases and the possible negative consequences decisions tainted with bias may have on others. This is particularly interesting when considering discipline disproportionality. Teachers are often faced with split-second decisions in their classroom concerning student behavior and discipline.

Implicit racial attitudes have been identified in children as young as three (Qian et al., 2016; Setoh et al., 2017). Interestingly, Qian et al. (2016) found implicit racial attitudes seemed to evolve over time. More to the point, Setoh et al. (2017) reported children not yet aware racial bias was socially undesirable had similar levels of explicit and implicit racial attitudes (while the adults did not). These findings indicate implicit racial attitudes develop early in life and change over time. Glock and Klaproth (2017) found teachers in Germany held implicit racial biases against Turkish students (an ethnic minority) while Hannon et al. (2013) found teachers preferred African American female students with lighter skin over African American female students with darker skin. Similarly, Okonofua and Eberhardt (2015) found teachers were more likely to hold lower regard for hypothetical students based on student name alone (e.g.–James vs. Jamaal).

Although the literature on implicit racial bias is growing, the mechanism to help individuals become aware of possible bias is questionable. Dovidio and Gardner (2000) believe educated professionals are well aware of the negative connotations surrounding explicit racial attitudes. Smolkowski et al. (2016) and McIntosh et al. (2017) purport teachers, in particular, are acutely aware of this phenomenon and will make every effort to appear racially non-biased in their classroom and community. Notwithstanding, Smolkowski et al. believe when teachers are unaware of their implicit bias and face split-second decisions in their classroom or school, the decision is likely tainted with bias. McIntosh et al. (2014) propose teachers may be able to reduce their bias if they are made aware of it and are given strategies to mitigate the negative effects bias may have on their discipline decisions.

Devine et al.'s (2012) intervention was designed to make individuals aware of their bias and gives them strategies to reduce or overcome their bias. The intervention includes a series of vignettes describing situations where bias may be present and provides individuals strategies designed to reduce or eliminate bias in a systematic manner. The intervention was found effective at reducing bias as measured by the Race IAT in undergraduate psychology students (Devine et al., 2012). The Race IAT is a latency based cognitive assessment that measures implicit bias through association (black skin/pleasant image vs. white skin/unpleasant image or vice versa). Pepis (2017) conducted a similar study using Devine et al.'s (2012) intervention with pre-service teachers. Although her study did not produce statistically significant results, participants were able to lower their bias as measured by the IAT over time.

Methodology

A sequential explanatory mixed methods design was used in this study. This design was employed to investigate if teacher awareness of implicit bias changed over time after being exposed to an intervention designed to reduce or eliminate bias and teacher perception of the process used to make them aware of their bias. Once the quantitative data were analyzed, interview questions were developed to better understand teacher perception of the study and explain the largely nonsignificant findings from the quantitative portion of the study. Creswell and Plano Clark (2017) state this design "can be used to explain the mechanisms through qualitative data that shed light on why the quantitative results occurred and how they might be explained" (p. 77).

For the quantitative methods of data collection, volunteer in-service teachers participated in three online Qualtrics surveys over three weeks. The specific data

collected and analyzed varied by the control and experimental group. Both the control group and experimental group took the Teacher Multicultural Awareness Survey (TMAS) as a pretest and posttest measure. The experimental group was asked to take the Race Implicit Association Test (IAT) at time points 1, 2, and 3. Three research questions were answered for the quantitative portion of this study. Research question 1 investigated if there was a significant difference between pretest and posttest scores on Ponterotto's (1995a) TMAS by the control group and experimental group. Research question 2 investigated if there was a significant difference among participants who scored low or high on the pretest TMAS on the final IAT score. Research question 3 investigated if there was a significant difference in participant repeated measure scores on the IAT.

The qualitative method of data collection included interview data collected from the nine teachers who participated in the quantitative portion of the study. Research question 4 was the cornerstone for the qualitative portion of this study: in what ways do the interview data of licensed teachers about their views on the importance of implicit bias in classrooms provide an explanation for any quantitative results from the IAT or TMAS? Creswell and Plano Clark (2017) describe this variation of the explanatory sequential design as the "follow-up explanations variant" (p. 82). Five participants from the experimental group and five participants from the control group indicated they were willing to be interviewed for the qualitative portion of this study. One participant from the control group was unable to participate.

A one-way analysis of covariance was used to answer research question 1, a one-way analysis of variance was used to answer research question 2, and a within-subject repeated-measures analysis of variance was used to answer research question 3. The

qualitative questions for research question 4 were designed to allow participants an opportunity to describe how their perspectives changed over time by participating in the study (Maxwell, 2013).

The demographic data gathered from the 60 participants who completed the pretest and posttest TMAS indicated most were female (57), White (55), and taught at the elementary level (37). Participants ranged in age from 25–55+, and reported varying years of experience. The education level of the participants included 13 with a bachelor's degree, 29 with a master's, and 18 who held an education specialist's degree.

Before taking Devine et al.'s (2012) intervention, 52 of the participants were concerned about equity in school discipline, 50 participants believed equity in school discipline affected the academic outcomes of students in their school, and 48 participants thought equity in school discipline affected the climate of their school. After completing the study, 45 participants were concerned with equity in school discipline, 48 participants believed equity in school discipline affected the academic outcomes of students in their school, and 44 participants thought equity in school discipline affected the climate of their school. Fifty-one participants indicated they thought participating in the study would influence their professional practice.

Fisher's exact tests were conducted to compare the participants before the pretest and after the posttest TMAS by the control group and experimental group. The groups were compared on their level of concern on equity in school discipline, if equity in school discipline affected the academic outcomes of their students, and if equity in school discipline affected the climate of their school. Only one Fisher's exact test was significant. After taking the posttest TMAS, participants in the control group were

significantly more concerned about the topic of equity in school discipline than the experimental group (p = .001, 95% CI [2.08, 117], OR = 11.2).

Results

Quantitative Findings

Research question 1 was designed to measure if teacher awareness of and sensitivity to multicultural issues in their classroom changed over time by the control group and experimental group. All participants took the Teacher Multicultural Attitude Survey (TMAS). Experimental group participants were also asked to complete the Race Implicit Association Test (IAT). The TMAS is a 20 item Likert-type instrument designed specifically for in-service classroom teachers. Scores can range from 20–100. A score of 20 would indicate a teacher has little or no appreciation and awareness of multicultural teaching issues in their classroom while a score of 100 would indicate a teacher is highly appreciative and aware of multicultural teaching issues in their classroom. A one-way analysis of covariance was used to answer research question 1. The overall mean pretest score for all 60 participants was M = 82 (SD = 7.32), higher than the overall mean posttest score for all 60 participants (M = 81, SD = 8.19). Control group participants (n = 34) had an overall mean pretest score of M = 83.88 (SD = 6.29) and overall mean posttest score of M = 82.74 (SD = 6.93). Experimental group participants (n = 26) had an overall mean pretest score of M = 79.54 (SD = 7.94) and overall mean posttest score of M = 78.73 (SD = 9.23). It is of interest the mean TMAS scores for all the participants, control group participants, and experimental group participants trended down from pretest to posttest. This indicates participants had less appreciation and awareness of multicultural teaching issues in their classroom after participating in the study. A one-way analysis of

covariance (ANCOVA) was then conducted to determine if there was a significant difference in posttest TMAS scores by the control and experimental group after controlling for pretest TMAS scores. After adjusting for pretest TMAS scores, there was not a significant difference in posttest TMAS scores between the control and experimental group ($F(1, 57) = 0.07$, $p = .80$).

Research question 2 investigated if participants who scored low or high on the pretest TMAS had significantly different IAT scores at time point 3 (IAT3). IAT scores are bound at the ±2 level. Race IAT scores with a positive value indicate an individual has a preference for white skin, and Race IAT scores with a negative value indicate a preference for black skin, respectively. This question was limited to the seven experimental group participants who completed IAT assessments at time point 1, 2, and 3. Demographic characteristics of the seven participants who met this requirement indicated six teachers were female, White, taught at the elementary level, and had over 15 years of experience. The teachers in this group included two with a bachelor's degree, two with a master's degree, and three with an education specialist's degree. The low group ($n = 4$) pretest TMAS mean was $M = 73$ and high group ($n = 3$) pretest TMAS mean was $M = 85.33$. The low group mean IAT3 score was $M = 0.14$ and high group mean IAT3 score was $M = -0.29$. This signifies teachers who scored low on the pretest TMAS had little to no preference for white skin at time point three, and teachers who scored high on the pretest TMAS had a slight preference for black skin at time point three. The one-way ANOVA, however, indicated there was not a significant difference on mean IAT3 scores among participants who scored low or high on the pretest TMAS , $F(1, 5)$, $p = .29$.

Research question 3 examined if participant IAT scores changed over time. A within-subject repeated-measures ANOVA was computed to investigate whether there was a significant difference in participant repeated measure IAT scores. Only the seven experimental group participants who completed all 3 IAT assessments were included in this analysis. The overall mean IAT1 score was $M = 0.42$, the overall mean IAT2 score was $M = 0.09$, and the overall mean IAT3 score was $M = -0.04$. The within-subject repeated-measures ANOVA indicated there was no significant difference in participant repeated measure IAT scores ($F(2, 12) = 2.23$, $p = .15$).

Although no significant difference was found in repeated measure IAT scores, it is worth noting scores did trend down over time. At time point 1, the mean IAT score was 0.42, indicating most participants had a moderate preference for white skin. At time point 2, the mean IAT score dropped to 0.09, indicating individuals had little to no preference for individuals with white skin. Remarkably, at time point 3, individuals had a mean IAT score of -0.04, indicating individuals had little to no preference for individuals with *black* skin. This suggested although the change in scores was not statistically significant, the underlying premise of implicit bias changing over time *did* occur. As measured by the IAT, participants started the study with a moderate pro-white skin bias and ended the study with little to no pro-black skin bias. This finding would suggest making concerned individuals aware of implicit bias and giving them strategies to reduce or eliminate their bias, was successful. Devine et al. (2012) agreed this change in preference is significant and lends credence to her belief individuals must be concerned about implicit bias and motivated to change before they can reduce or eliminate their bias. Interestingly, participant *appreciation and awareness* about

multicultural teaching issues in the classroom, as measured by TMAS mean scores, decreased for both the control group and experimental group participants in the study.

Qualitative Findings

Both control and experimental group participants were included in the interview portion of this study. Interviewees were able to view the questions and review the results of the quantitative portion of the study before being interviewed. The questions centered on teacher perception of the surveys and Devine's (2012) intervention, the differential dropout rate between the control and experimental groups, the change in preference from white skin to black skin as measured by the Race IAT, significance of the study concerning equity in school discipline, teacher effect on students, personal reflections concerning implicit bias, the role of K-12 leadership in addressing implicit bias, and thoughts on if training in this area would affect the teaching and learning at their school.

Interview responses on what was most important to control group and experimental group participants revealed two distinct themes. Experimental group participants all reported taking the Race IAT was the one aspect of the study they remembered most. This reaction was somewhat expected according to research indicating the results from the Race IAT may cause participants to experience powerful emotions of distress, discomfort, disregard, disbelief, acceptance, or react defensively (Clark & Zygmunt, 2014; Howell et al., 2014). Several respondents admitted the IAT surprised them, explaining they were not expecting the images and words to switch in the middle of the IAT. Olivia confessed when the words and pictures "switched", it took a minute for her mind to adjust. Similarly, Aurora mentioned the switching of words and pictures and stated plainly, "you really have to think." Throughout the interview process

these individuals repeatedly brought up the IAT–even when they were asked about the TMAS. Ava stated she thought "individuals were going to change some" after taking the IAT. Experimental group participants also talked about the emotions they felt after taking the IAT. Amelia stated she "felt bad about myself thinking that I might have some biases" and Olivia "felt angry at them, the buttons I had to push when the words and pictures switched" and thought others might have felt "called out" once they received their IAT results. Ava, Amelia, and Olivia all stated the IAT made them "stop and think" about how implicit bias may affect their decision making in the classroom. All five respondents felt the IAT made them more aware of implicit bias. Ava remarked "I'd like to take it (IAT) again."

Control group participants, on the other hand, hardly mentioned the IAT when reflecting on what they remembered most about the surveys. Once John figured out he was not going to take the IAT he "was kind of curious as to how I would have scored." Participants from the control group thought Devine's intervention was effective at increasing their awareness of implicit bias. The intervention did seem to increase participant's awareness of bias and the overall concern individuals had on how these biases may negatively affect their students (or society at large). James mentioned his overall "self-awareness" increased and Mary expressed one of her "biggest fears (as a teacher) is that I would have some . . . hidden racism or implicit bias and not realize it." Jennifer admitted she found "there were biases I didn't even realize I had" and considered the intervention "eye-opening." Overall, the respondents in the control group expressed feelings of curiosity and a desire to improve as driving factors to participate in the study and considered the intervention and experience of participating professionally beneficial.

Mary stated she hoped participating would make her a better teacher and this would be something "that would be beneficial to my faults."

Experimental group participants were split on why they thought the control group mean TMAS scores were higher than the experimental group mean scores. Charlotte and Ava were not surprised the control group mean TMAS scores were higher than their group. Charlotte figured most professionals can "control what they are thinking" and will usually say the right thing to appear professional or non-racist. Although she did not explicitly mention the IAT or the anonymous nature of the surveys, her response indicated she thought the scores were lower for her group because of the IAT. She considered herself colorblind, stating she does not see "color and races" she sees "children" in her classroom. Ava also thought the scores were different because of the IAT. She mentioned the IAT did not affect her as much as other teachers, because as an ESOL (English to Speakers of Other Languages) teacher, she considered herself culturally competent. Amelia, Olivia, and Aurora, however, were perplexed by the difference in group scores. They figured TMAS scores would have gone up for the experimental group *because* of the IAT. Amelia put it plainly: "I would have thought the experimental group (TMAS score) would have gone up and made the overall score higher for the experimental group." Amelia considered the groups may have just been made up differently and one group was more culturally sensitive than the other. Aurora thought about it for a moment, and after realizing the control group did not take the IAT, believed control group participants were not as honest on the TMAS as experimental group participants and "were basically assessing themselves."

Nearly all control group participants explicitly stated they thought the difference in group mean TMAS scores was due to the IAT. Mary thought the IAT probably affected experimental group participants more than usual due to the ongoing civil unrest the United States was experiencing during the summer of 2020. Jennifer thought the control group participants were not as honest as experimental group participants and answered TMAS questions how they felt they *should* answer rather than answer honestly. John also figured the IAT had something to do with the difference in scores and predicted experimental group participants were probably "more aware" of their personal biases. Interestingly, both groups mean TMAS scores dropped slightly. TMAS scores are considered to measure teachers' explicit attitudes about multicultural issues in the classroom. This indicates participants rated their explicit racial attitudes less favorably after participating in the study. Devine et al. (2012) theorized the intervention would cause participants to become more culturally aware and care more about the unintended consequences implicit bias causes for affected groups. This data suggests as awareness and knowledge is gained by utilizing the intervention, individuals realize they may not be as culturally neutral as they thought.

Considerably more experimental group participants dropped out of the study either partly or all together than did control group participants. Both groups were split on if they considered the dropout rate reasonable or explainable. Experimental group participants Charlotte and Aurora considered the dropout rate explainable but not necessarily reasonable. Charlotte figured some participants may have experienced technical difficulties with the IAT or surveys (or both) and simply quit the study. Aurora considered the extra time experimental group participants were inherently faced with by

taking the IAT at multiple time points was too much for some and "people got tired of doing the test (IAT)." Control group participants Mary and John considered the dropout rate both reasonable and explainable. Mary thought the IAT was likely personal for some and they might not have been ready for the assessment, "people maybe were [*sic*] scared to find out more about themselves." John thought experimental group participants may have tired from taking the IAT, but, "in the back of my mind, if people weren't happy with their (IAT) score, maybe it said they had a strong preference for white skin, they may not want to do it again because it was right." This data strongly correlates with the findings of both Clark and Zygmunt (2014) and Howell et al. (2014). They found if individuals were not prepared to receive their IAT results or if IAT results were extremely different than their self-reported levels of explicit bias, they would most likely disengage from the process or become defensive and more entrenched in their previous beliefs (Clark & Zygmunt, 2014; Howell et al., 2014).

Experimental group participants Ava, Amelia, and Aurora and control group participants James and Jennifer did not think the dropout rate difference was reasonable or explainable. They did not consider dropping out of the study and were curious why some participants would finish the surveys but not the IAT. Amelia did consider the IAT "frustrating" when her scores were different than what she wanted them to be and Olivia thought some participants did not like taking the IAT multiple times. James thought most teachers should be professionally willing to grapple with issues like implicit bias. In a recent study, however, Starck, Riddle, Sinclair, and Warikoo (2020) found teachers were like other professionals in both implicit and explicit bias measures. Nearly all professionals in the two national data sets they utilized in their study held some degree of

pro-white implicit and explicit racial attitudes (Starck et al., 2020). These findings, Starck et al. suggest, should encourage schools and teachers to investigate ways to help teachers reduce or eliminate either bias in general or the effects of bias on students in particular (2020).

Nearly all teachers from both groups attributed the change in preference from white skin to black skin due to teacher effort. More precisely, they believe the teachers who completed the IAT three times were trying to drop their score. The rationale, however, varied among the respondents. Experimental group participants Ava and Amelia considered it a professional responsibility to lower their IAT scores. Amelia thought "everybody (teachers) is trying to be more racially sensitive" and Ava thought it part of a teachers "job to learn other people's culture." Olivia and Aurora considered it a cognitive challenge to lower their scores. Olivia stated the second and third time she took the IAT, "I was prepared for the changes (key switches for black/white and pleasant/unpleasant)." Aurora thought the IAT made her "sit there and think", she knew what the assessment was trying to measure and how it was measuring it. Aurora thought teachers might be given IAT assessments other than race in the future then hesitated: "but, then, you want them to be more conscious of it (racial implicit bias)." Control group participants James and Jennifer thought teachers in the experimental group intentionally tried to lower their scores. James thought teachers stopped seeing color and "focused on the keys" and Jennifer posited the inherent experience of taking the IAT and surveys could "push you (to change)." She thought as individuals became aware of their bias and acknowledge it, true change would manifest. Control group participants Mary

and John did not weigh in on this change. Mary studied the Race IAT to understand how it measures implicit bias and questioned the validity of the instrument.

Nearly all participants from both the control group and experimental group considered the experience of participating in the study personally significant and affected their views on the topic of equity in school discipline. Charlotte, from the experimental group, provided the only disconfirming response. She recently completed a yearlong professional learning module and did not think this study had nearly the effect on her. Interestingly, Gregory, Hafen, Ruzek, Mikami, Allen, and Pianta (2016) found teachers who participated in an intensive yearlong professional development program (MTP-S) centered on student teacher interactions were able to close the achievement gap for African American students, were less likely to issue ODR's to African American students, and were less likely to issue subsequent ODR's to African American students (if applicable). Charlotte's belief a three-week study is not comparable to a yearlong process is worth noting. All other control and experimental group participants indicated the process increased their awareness of implicit bias. Experimental group participant Ava thought the topic should be addressed regularly in K-12 settings. Amelia, Olivia, and Aurora thought participating in the study would help them slow down and think about the possible negative effects bias may have on their classroom decisions. All four control group members thought participating in the study increased their level of awareness concerning implicit bias in their classroom. Mary asserted she is constantly looking for ways to improve her efficacy. She claimed if examining her practice and beliefs would help her improve and reach more of her students, she would do it. Jennifer also conveyed her desire to improve "and do whatever I can to help my students." In a

similar vein, John strives to be "more aware how can [*sic*] I be more helpful or more aware of what my kids need."

Charlotte, from the experimental group, once more provided the only disconfirming answer when compared to all other control and experimental group participants on the topic of multicultural teaching issues effect on students. She believes too much emphasis has been placed on the topic of race in our society. She sees "a lot of problems that has [*sic*] absolutely nothing to with race." It is worth noting the intense, year-long program Gregory et al. (2016) found effective at reducing the discipline gap and closing the achievement gap does not "explicitly focus on raising teacher consciousness about implicit bias or institutional racism . . . rather, it focuses on skills in effectively interacting with any student" (p. 186). Experimental group participants Ava, Amelia, Olivia, and Aurora thought students were affected in several ways by the issues of race, culture, and bias (Blitz et al., 2016). Ava and Oliva believe students should be taught to be more tolerant of their peers and Amelia suggested many of the issues are taught and learned in environments outside the classroom. All four control group participants thought the issues of race, culture, and bias had a direct negative impact on their students. James mentioned the gang culture prevalent in a neighborhood near his elementary school while Mary and Jennifer believe lower teacher expectations and subsequent academic disengagement many African American students face create a culture of failure. John, who teaches at the high school level, thought the overwhelming amount of news coverage over the summer of 2020 would cause cultural issues at his school. The Black Lives Matter movement and the never-ending news cycle of White police officers shooting unarmed Black men, John continued, was bound to have an

impact–especially when African American students look around their school and all they see "are a bunch of White teachers."

Nearly all teachers were able to recall and share a personal story concerning implicit bias in their school or classroom. Experimental group participants Charlotte and Aurora were not able to recall a personal experience. All four control group participants were able to recall and share a story concerning implicit bias. Ava, Amelia, and Olivia's stories were deeply personal and affect them still. Ava shared that her relationship with an African American student blossomed after she slowed down and listened to him. She confessed the memory influences her classroom practice daily. Amelia and Olivia still worry about how their decisions may have had long term consequences or negative effects for the students they described. Control group participant Jennifer spoke about her first year as a teacher in a school with predominantly African American students. James, Mary, and John all spoke about the importance of building relationships with the students in their classrooms.

Several respondents from the experimental group and all control group participants thought K-12 leaders should take the time to build honest and open professional relationships with their teachers so meaningful conversations could occur. Charlotte and Ava from the experimental group were not asked this question. Amelia thought before teachers are comfortable thinking about implicit bias, they need to be comfortable talking about it with their building administrator. Aurora agreed, adding teachers and leaders need to have conversations about the topic. As an example, Aurora mentioned the civil unrest in the United States during the summer of 2020 stating, "everything going on in the nation, you know, you feel bad sometimes for being White."

James put the onus on leaders to become cognizant of the effects implicit bias may have on the students at their school and model the correct behavior for teachers. He believes leaders need "to make it a point" and "make it a practice" for the teachers in their school. Mary thought clear expectations concerning behavior and discipline from school leaders would alleviate many issues stating "my God–don't sweat the small stuff!"

The teachers had varying opinions on how to implement effective training on the topic of implicit bias at their school. Experimental group participants Charlotte and Amelia thought the training would need to be voluntary and pragmatic. Charlotte suggested teachers would benefit from visiting the homes and communities of their African American students. Amelia asserted mandatory teacher training is already overwhelming, but considered it might be helpful for teachers who "really want to improve." Ava, Olivia, and Aurora contend implicit bias professional development should be a priority for all K-12 educators. Ava clarified, however, if teachers were able to attend training, they needed to be given adequate time to reflect on and practice what they learn. Olivia admitted she was not certain how to address the issue, but thought it needs to be addressed. Mary, Jennifer, and John from the control group were convinced training aimed at raising awareness on multicultural issues in the classroom would benefit the teaching and learning in their school and classroom. Jennifer speculated as teachers become comfortable with and aware of implicit bias, effective analysis and problem solving could occur.

Limitations and Assumptions

The researcher has worked at the elementary, middle, and high school level as a teacher. All of the schools were Title I schools with high rates of poverty and ethnically

diverse student populations. The researcher was also district IEP coordinator for one year. The researcher acknowledges a certain amount of bias could exist, especially during the qualitative portion of this study. The researcher acknowledges all teachers in this study were employed at a Title I school the year before participating in the study. It should be noted teacher efficacy or perceived efficacy were not examined in this study.

The researcher acknowledges the small sample size of participants in the quantitative portion of this study were lower than expected and greatly limit the generalizability of the results. In addition, teachers interested in the topic of equity in school discipline were recruited to participate, further limiting the generalizability of the study. The researcher also acknowledges teachers in the study were not able to complete the 2019-2020 school year due to the novel coronavirus and subsequent closing of all public schools in the state. Moreover, the civil unrest surrounding the murder of George Floyd and subsequent racial riots in cities across the nation occurred during the quantitative data collection portion of this study.

The quantitative portion of the study ended in June of 2020. The qualitative portion of the study, utilizing participants from the quantitative portion of the study, did not commence until August of 2020. The researcher acknowledges this gap in data collection is another limitation of the study. Although typical for studies utilizing the explanatory sequential design, the timing could have been shortened so respondent memory of the surveys and IAT were clear.

Suggestions for Future Research

Interview data from both the control and experimental group illustrate the power of the Race IAT. Future research could include different types of the IAT (politics,

gender, religion) with Devine et al.'s (2012) intervention and Ponterotto's (1995a) TMAS to explore the racial construct in a less threatening manner. Several participants also mentioned how becoming aware of implicit bias would improve their teacher efficacy. Therefore, it is suggested future iterations of this study examine student and teacher perception of efficacy as a pretest and posttest measure.

In the same way, future research could explore how K-12 leaders concerned about the topic of equity in school discipline utilizing a similar conceptual framework are effected. Charlotte mentioned how she considered the intensive year-long professional development course on implicit bias more meaningful and effective than this study. Therefore, future qualitative research could embed elements of this study within a program like MTP-S by combining explicit instruction on the racial construct within a teacher coaching model designed to improve student/teacher interactions in the classroom. Finally, future research could replicate this study in other settings and locations with a larger teacher participant pool to investigate the generalizability of the results for teachers concerned about the topic of equity in school discipline.

Conclusion

This study includes some unique characteristics relatively new to the long-known problem of discipline disproportionality. By utilizing in-service, volunteer teachers concerned about the topic of equity in school discipline and providing teachers with an intervention designed to reduce or eliminate their bias, using an instrument specifically designed to measure teacher explicit racial attitudes, and the Race IAT for implicit racial bias, this study was able to build on previous research conducted by Devine et al. (2012) and Pepis (2017). By utilizing a control and experimental group, two dependent

variables measured teacher change over time on the construct of prejudice. Additionally, nine teachers who participated in the quantitative portion of the study were interviewed to help understand the quantitative findings. Only in-service classroom teachers were eligible to participate.

Research question 1 investigated if teacher appreciation and awareness of multicultural issues in the classroom as measured by the TMAS were significantly different by the control and experimental group. Research question 2 measured if experimental group participants who scored low or high on the pretest TMAS had significantly different IAT scores at time point 3. Research question 3 measured if teacher repeated measure IAT scores changed over time.

Although the literature on discipline disproportionality and equity in school discipline is substantial, the effect implicit bias may have on teacher decisions at the classroom level is yet emerging. This study has major implications for schools and districts seeking to improve in this area. Teachers seeking to improve their pedagogical efficacy seem willing to grapple with this socially sensitive topic when given the time and freedom to do so. If K-12 leaders address this topic through the lens of improved teacher efficacy, more teachers may be willing to examine racial disparities in their classroom and school. Although the quantitative findings of the study were not statistically significant, the change in teacher preference for the seven who completed the study for the experimental group from a moderate preference for white skin to little to no preference for black skin is intriguing. Coupled with the overall decrease in mean TMAS scores, the findings are nearly perplexing. The teachers who participated in this study were all on paid "leave" due to SARS-CoV-2 (severe acute respiratory syndrome

coronavirus 2). They had the time and resources to participate. The differential dropout rate when comparing the control and experimental group and the extreme dropout rate of experimental group participants taking the IAT was a cautionary development.

The interview data made clear how powerful the Race IAT is. Experimental group participants referenced the IAT repeatedly, while hardly mentioning Devine et al.'s (2012) intervention. In contrast, control group participants remembered Devine's intervention clearly and spoke on how it changed their awareness of implicit bias. Yet– the power of the IAT seemed to come at a cost. It is worth noting the experimental group mean TMAS pretest scores (M =79.54), taken before the initial IAT, were lower by any metric than the overall control group mean TMAS pretest scores (M = 83.88).

The interview data from respondents indicated their overall awareness and acknowledgment of implicit bias increased by participating. Charlotte notwithstanding, participants felt the process was a good use of their time professionally. Most striking, however, was the overwhelming desire to improve. Every teacher mentioned how important improving their teacher efficacy was to them professionally. Control group participant James summed it up:

> Discipline disproportionality (leads to) lower reading, lower Lexile levels, lower achievement, higher poverty, higher teen pregnancy, higher mortality, the list goes on. And I could correlate every one of those findings with discipline. It is not about being scared. It's about making the change. So at some point the change has to be made, otherwise the issue will never go away.

Although this study failed to show a statistically significant drop in IAT scores over time, and teacher appreciation and awareness of multicultural issues in the

classroom as measured by the TMAS dropped from pretest to posttest, it did give teachers concerned about the topic of equity in school a discipline a voice in this long raging debate. The qualitative data from this study indicate teachers seeking to improve will go to considerable effort to increase their efficacy. The fact teachers who completed the Race IAT 3 times were able to change from a moderate preference for white skin to little to no preference for black skin is remarkable. Even if teachers were trying to lower their scores for personal gratification, they were motivated to change their thought patterns. Finding a pragmatic solution to the problem of equity in school discipline is not likely to come in the form of a standalone, one-time intervention. More than likely, a combination of coaching teachers coupled with careful instruction on the topic of implicit bias will prove effective at reducing or eliminating inequity in school discipline. It will not be easy. If it were, the problem would have been solved long ago.

References

Allen, Q. (2015). "I'm trying to get my A": Black male achievers talk about race, school and achievement. *The Urban Review, 47*(1), 209-231. doi:http://dx.doi.org/10.1007/s11256-014-0315-4

Amin, A. (2017). Implicit bias in the courtroom and the need for reform. *Geo. J. Legal Ethics, 30*, 575.

Anderson, K. A. (2015). An introduction to optimal resource theory: A framework for enhancing student achievement. *The Journal of Negro Education*, *84*(1), 25-39.

Arcia, E. (2006). Achievement and enrollment status of suspended students: Outcomes in a large, multicultural school district. *Education & Urban Society, 38*(3), 359-369. doi:10.1177/0013124506286947

Ary, D., Jacobs, L. C., Sorenson, C. K., & Walker, D. A. (2014). *Introduction to research in education* (9th ed.). Belmont, CA: Wadsworth.

Balfanz, R., Byrnes, v., & Fox, J. (2014). Sent home and put off-track: The antecedents, disproportionalities, and consequences of being suspended in the ninth grade. *Journal of Applied Research on Children, 5*(2), 1-19.

Beck, A. N., & Muschkin, C. G. (2012). The enduring impact of race: Understanding disparities in student disciplinary infractions and achievement. *Sociological Perspectives, 55*(4), 637-662. doi:http://dx.doi.org/10.1525/sop.2012.55.4.637

Behnken, M. P. (2014). Linking ADHD to incarceration among African Americans. *The ADHD Report, 22*(7), 9-16. doi:http://dx.doi.org/101521adhd20142279

Bell, E. E. (2015). Understanding African American males' schooling experiences: A qualitative inquiry. *The Qualitative Report, 20*(8), 1260-1269. doi:10.1177/0042085903256220

Bleyaert, B. (2009). *Disproportional high school suspension rates by race and ethnicity. Research brief*. Retrieved from http://library.valdosta.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED537954&site=ehost-live

Blitz, L. V., Anderson, E. M., & Saastamoinen, M. (2016). Assessing perceptions of culture and trauma in an elementary school: Informing a model for culturally responsive trauma-informed schools. *The Urban Review, 48*(4), 520-542. doi:http://dx.doi.org/10.1007/s11256-016-0366-9

Boneshefski, M. J., & Runge, T. J. (2014). Addressing disproportionate discipline practices within a school-wide positive behavioral interventions and supports framework: A practical guide for calculating and using disproportionality rates. *Journal of Positive Behavior Interventions, 16*(3), 149. doi: 10.1177/1098300713484064

Bottiani, J. H., Bradshaw, C. P., & Mendelson, T. (2017). A multilevel examination of racial disparities in high school discipline: Black and white adolescents' perceived equity, school belonging, and adjustment problems. *Journal of Educational Psychology*, *109*(4), 532. doi: 10.1037/edu0000155

Brown, K. E., & Steele, A. S. L. (2015). Racial discipline disproportionality in montessori and traditional public schools: A comparative study using the relative

rate index. *Journal of Montessori Research, 1*(1), 14. doi:10.17161/jomr.v1i1.4941

Brown v. Board of Education, 347 U.S. 483 (1954).

Children's Defense Fund (1975). *School suspensions: Are they helping children?* Cambridge, MA. Retrieved from http://diglib.lib.utk.edu/cdf/data/0116_000050_000205/0116_000050_000205.pdf

Clark, P., & Zygmunt, E. (2014). A close encounter with personal bias: Pedagogical implications for teacher education. *The Journal of Negro Education, 83*(2), 147-161.

Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Los Angeles, CA: Sage publications.

Devine, P. G. (2016). *Outline of new intervention*. Unpublished instrument.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of experimental social psychology, 48*(6), 1267-1278. doi:10.1016/j.jesp.2012.06.003

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive racism and selection decisions: 1989 and 1999. *Psychological Science (0956-7976), 11*(4), 315.

Eastman, J. T. (2015). The wild (white) ones: Comparing frames of white and black deviance. *Contemporary Justice Review, 18*(2), 231-247. doi:10.1080/10282580.2015.1025634

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology, 54*, 297-327. doi:10.1146/annurev.psych.54.101601.145225

Fiester, H. R., & Gibson, N. (2015). The effect of office discipline referrals, race, gender, and beginning of the year fluency scores on reading comprehension for fifth grade students. *International Journal of Arts and Commerce,* 4(6), 45-58.

Fisher, J. D. (2011). *School composition and disproportionality in discipline: A school-level analysis of disciplinary actions against various offense types in the United States, 2007–2008* (Doctoral Dissertation). Retrieved from UMI. (3450545)

Forscher, P. S. (2016). *Method*. Retrieved from osf.io/gv27n

Forscher, P. S., Mitamura, C., Dix, E. L., Cox, W. T., & Devine, P. G. (2017). Breaking the prejudice habit: Mechanisms, timecourse, and longevity. *Journal of Experimental Social Psychology*, *72*, 133-146.

Fox, J. (2016) *Applied Regression Analysis and Generalized Linear Models*, (3rd ed.). Thousand Oaks, CA: Sage.

Freeman, K. J., & Steidl, C. R. (2016). Distribution, composition and exclusion: How school segregation impacts racist disciplinary patterns. *Race and Social Problems, 8*(2), 171-185. doi: http://dx.doi.org/10.1007/s12552-016-9174-9

French, R., & Ernsthausen, J. (2014, January 19). Race sways severity of school discipline: Black students are suspended much more often than whites. State rates even worse than national figures; new approach urged. The Atlanta Journal Constitution. Retrieved from http://www.myajc.com

Gass, K. M., & Laughter, J. C. (2015). "Can I make any difference?" Gang affiliation, the school-to-prison pipeline, and implications for teachers. *The Journal of Negro Education, 84*(3), 333-347.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological bulletin*, *132*(5), 692. doi: 10.1037/0033-2909.132.5.692

Glock, S., & Klapproth, F. (2017). Bad boys, good girls? Implicit and explicit attitudes toward ethnic minority students among elementary and secondary school teachers. *Studies in Educational Evaluation, 53*, 77-86. doi:https://doi.org/10.1016/j.stueduc.2017.04.002

Gordon, R., Piana, L. D., & Keleher, T. (2000). *Facing the consequences: An examination of racial discrimination in U.S. public schools*. Retrieved from http://library.valdosta.edu:2048/login?url=http://search.proquest.com/docview/62339221?accountid=14800

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4-27. doi:10.1037/0033-295X.102.1.4

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology, 108*(4), 553-561. doi:10.1037/pspa0000016

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480.

Greenwald, A. G., Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie, 48*(2), 85–93. https://doi.org/10.1026//0949-3946.48.2.85

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197-216. doi:10.1037/0022-3514.85.2.197

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17. doi: 10.1037/a0015575

Gregory, A., Allen, J. P., Mikami, A. Y., Hafen, C. A., & Pianta, R. (2014). Eliminating the racial disparity in classroom exclusionary discipline. *Journal of Applied Research on Children, 5*(2), Article 12.

Gregory, A., Hafen, C. A., Ruzek, E., Mikami, A. Y., Allen, J. P., Pianta, R. C. (2016). Closing the racial discipline gap in classrooms by changing teacher practice. *School Psychology Review, 45*(2), 171-191.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin?*, Educational Researcher* 39(1), 59-68. doi: 10.3102/0013189X09357621

Hannon, L., Defina, R., & Bruch, S. (2013). The relationship between skin tone and school suspension for African Americans. *Race and Social Problems, 5*(4), 281-295. doi:http://dx.doi.org/10.1007/s12552-013-9104-z

Hartmann, D., Croll, P. R., Larson, R., Gerteis, J., & Manning, A. (2017). Colorblindness as identity: Key determinants, relations to ideology, and implications for attitudes about race and policy. *Sociological Perspectives, 60*(5), 866-888. doi:10.1177/0731121417719694

Heilbrun, A., Cornell, D., & Lovegrove, P. (2015). Principal attitudes regarding zero tolerance and racial disparities in school suspensions. *Psychology in the Schools, 52*(5), 489-499. doi:http://dx.doi.org/10.1002/pits.21838

Howell, J. L., Gaither, S. E., & Ratliff, K. A. (2014). Caught in the middle. *Social Psychological and Personality Science, 6*(4), 373-381. doi:10.1177/1948550614561127

Hughes, J., & Kwok, O. (2007). Influence of student-teacher and parent-teacher relationships on lower achieving readers' engagement and achievement in the primary grades. *Journal of Educational Psychology, 99*(1), 39-51. doi:10.1037/0022-0663.99.1.39

Indiana University (2016). *The equity project: Understanding equity*. Retrieved from http://www.indiana.edu/~equityiu/understanding-equity/

Irby, D. J. (2013). Net-deepening of school discipline. *The Urban Review, 45*(2), 197-219. doi:http://dx.doi.org/10.1007/s11256-012-0217-2

Krezmien, M. P., Leone, P. E., & Achilles, G. M. (2006). Suspension, race, and disability: Analysis of statewide practices and reporting. *Journal of Emotional and Behavioral Disorders, 14*(4), 217-226.

Kubota, J. T., Peiso, J., Marcum, K., & Cloutier, J. (2017). Intergroup contact throughout the lifespan modulates implicit racial biases across perceivers' racial group. *PLOS ONE, 12*(7): e0180440. doi:10.1371/journal.pone.0180440

Losen, D., Hewitt, D., & Toldson, I. (2014). Eliminating excessive and unfair exclusionary discipline in schools: Policy recommendations for reducing disparities. *Discipline Disparities: A Research-to-Practice Collaborative at Indiana University*. Retrieved from https://www.njjn.org/uploads/digital-library/OSF_Discipline-Disparities_Disparity_Policy_3.18.14.pdf

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Los Angeles, CA: Sage Publications.

McIntosh, K., Ellwood, K., McCall, L., & Girvan, E. J. (2017). Using discipline data to enhance equity in school discipline. *Intervention in School and Clinic, 53*(3), 146-152. doi:10.1177/1053451217702130

McIntosh, K., Girvan, E. J., Horner, R., & Smolkowski, K. (2014). Education not incarceration: A conceptual model for reducing racial and ethnic disproportionality in school discipline. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 5*(2), 4.

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, *46*(3), 235-261.

Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science, 159*(3810), 56-63

Monahan, K. C., Vanderhei, S., Bechtold, J., & Cauffman, E. (2014). From the school yard to the squad car: School discipline, truancy, and arrest. *Journal of Youth and Adolescence, 43*(7), 1110-1122. doi:http://dx.doi.org/10.1007/s10964-014-0103-1

Mortenson, B. C. (2018). *The role of teacher implicit bias in the racial achievement gap* (Master's thesis). Retrieved from ProQuest. (10791341)

Nance, J. P. (2016). Dismantling the school-to-prison pipeline: Tools for change. *Arizona State Law Journal, 48*(2), 313-372.

Nance, J. P. (2017). Student surveillance, racial inequalities, and implicit racial bias. *Emory Law Journal, 66*(4), 765-837.

Nicholson-Crotty, S., Birchmeier, Z., & Valentine, D. (2009). Exploring the impact of school discipline on racial disproportion in the juvenile justice system. *Social Science Quarterly, 90*(4), 1003-1018.

Noltemeyer, A. L., Ward, R. M., & McLoughlin, C. (2015). Relationship between school suspension and student outcomes: A meta-analysis. *School Psychology Review, 44*(2), 224-240.

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2006). *The Implicit Association Test at age 7: A methodological and conceptual review*. In J. A. Bargh (Ed.), Automatic processes in social thinking and behavior (pp. 265 – 292). Hove, UK: Psychology Press.

Nosek, B. A., Smyth, F. L., Hanson, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., … Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology, 18,* 36-88. doi: 10.1080/10463280701489053

Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science, 26*(5), 617-624. doi:10.1177/0956797615570365

Onwuegbuzie, A. J., & Collins, K. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report, 12*(2), 281-316. Retrieved from https://nsuworks.nova.edu/tqr/vol12/iss2/9

Onwuegbuzie, A. J., & Leech, N. L. (2004). Post hoc power: A concept whose time has come. *Understanding Statistics*, *3*(4), 201–230. https://doi.org/10.1207/s15328031us0304_1

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*(2), 171–192. https://doi.org/10.1037/a0032734.supp (Supplemental)

Pearman II, F. A., Curran, C. F., Fisher, B., & Gardella, J. (2019). Are achivement gaps related to discipline gaps? Evidence from national data. *AERA Open*. https://doi.org/10.1177/2332858419875440

Pepis, T. (2017). *Implicit bias in pre-service teachers: A mixed methods approach* (Doctoral Dissertation). Retrieved from http://scholarworks.umass.edu/dissertations_2. (959)

Peterson, E. R., Rubie-Davies, C., Osborne, D., & Sibley, C. (2016). Teachers' explicit expectations and implicit prejudiced attitudes to educational achievement: Relations with student achievement and the ethnic achievement gap. *Learning & Instruction, 42*, 123-140. doi:10.1016/j.learninstruc.2016.01.010

Phinney, J. (1992). The Multigroup ethnic identity measure: A new scale for use with adolescents and young adults from diverse groups. *Journal of Adolescent Research, 7*, 156-176.

Plant, A., & Devine, P. (2009). The active control of prejudice: Unpacking the intentions guiding control efforts. *Journal of Personality and Social Psychology, 96*(3), 640-652.

Ponterotto, J. G. (1995a). Teacher multicultural attitude survey. New York, NY; Fordham University.

Ponterotto, J. G. (1995b). *The quick discrimination index*. New York, NY; Fordham University.

Ponterotto, J. G., Baluch, S., Greig, T., & Rivera, L. (1998). Development and initial validation of the teacher multicultural attitude survey (TMAS). *Educational and Psychological Measurement, 58*(6), 1002-1016.

Project Implicit (2011). Race IAT [Computer program]. Cambridge, MA; Harvard University.

Qian, M. K., Heyman, G. D., Quinn, P. C., Messi, F. A., Fu, G., & Lee, K. (2016). Implicit racial biases in preschool children and adults from Asia and Africa. *Child Development, 87*(1), 285-296. doi:10.1111/cdev.12442

Rezaei, A. R. (2011). Validity and reliability of the IAT: Measuring gender and ethnic stereotypes. *Computer in Human Behavior, 27*, 1937-1941. doi:10.1016/j.chb.2011.04.018

Seidman, I. (2012). *Interviewing as qualitative research: A guide for researchers in education and the social sciences* (4th ed.). New York: Teachers College Press.

Setoh, P., Lee, K. J., Zhang, L., Qian, M. K., Quinn, P. C., Heyman, G. D., & Lee, K. (2017). Racial categorization predicts implicit racial bias in preschool children. *Child Development*, 1-18. doi:10.1111/cdev.12851

Shollenberger, T. L. (2015). Racial disparities in school suspension and subsequent outcomes: Evidence from the National Longitudinal Survey of Youth. *Closing the school discipline gap: Equitable remedies for excessive exclusion*, 31-44.

Simmons, L. (2015). Profiting from punishment: Public education and the school security market. *Social Justice, 41*(4), 81-95.

Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. (2000). *The color of discipline: Sources of racial and gender disproportionality in school punishment*. Policy Research Report. Retrieved from http://library.valdosta.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED468512&site=eds-live&scope=site

Skiba, R., & Peterson, R. (2003). Teaching the social curriculum: School discipline as instruction. *Preventing School Failure: Alternative Education for Children and Youth*, *47*(2), 66-73.

Smolkowski, K., Girvan, E. J., McIntosh, K., Nese, R. N. T., & Horner, R. H. (2016). Vulnerable decision points for disproportionate office discipline referrals: Comparisons of discipline for African American and White elementary school students. *Behavioral Disorders, 41*(4), 178-195.

Starck, J. G., Riddle, T., Sinclair, S., & Warikoo, N. (2020). Teachers are people too: Examining the racial bias of teachers compared to other American adults. *Educational Researcher, 49*(4), 273-284. doi: 10.3102/0013189X20912758

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A. M. B., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: Implications for clinical practice. *American Psychologist, 62*(4), 271-286. doi:10.1037/0003-066X.62.4.271

Sullivan, A. L., & Bal, A. (2013). Disproportionality in special education: Effects of individual and school variables on disability risk. *Exceptional Children, 79*(4), 475-494.

Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, *99*(2), 253. doi: 10.1037/0022-0663.99.2.253

U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. (2020). *The condition of education: Characteristics of public school teachers.* Retrieved from: https://nces.ed.gov/programs/coe/indicator_clr.asp

U.S. Department of Education. (2016a). *Office for Civil Rights. 2013-2014 Civil Rights Data Collection: A first look*. Retrieved from https://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-look.pdf

U.S. Department of Education. (2016b). *School climate and discipline: Know the data.* Retrieved from http://www2.ed.gov/policy/gen/guid/school-discipline/data.html#featured

U.S. Department of Education (2016c). Office of Special Education and Rehabilitation Services. *Racial and ethnic disparities in special education: A multi-year disproportionality analysis by state, analysis category, and race/ethnicity.* Retrieved from https://www2.ed.gov/programs/osepidea/618-data/LEA-racial-ethnic-disparities-tables/disproportionality-analysis-by-state-analysis-category.pdf

U.S. Department of Justice, Civil Rights Division and the U.S. Department of Education, Office of Civil Rights (2014). *Dear colleague letter: Nondiscriminatory administration of school discipline*. Retrieved from http://www.justice.gov/crt/about/edu/documents/dcl.pdf.

Van den Bergh, L., Denessen, E., Hornstra, L., Voeten, M., & Holland, R. W. (2010). The implicit prejudiced attitudes of teachers: Relations to teacher expectations and the ethnic achievement gap. *American Educational Research Journal*, *47*(2), 497-527. doi: 10.3102/0002831209353594

Vanderhaar, J., Munoz, M., & Petrosko, J. (2014). Reconsidering the alternatives: The relationship between suspension, disciplinary alternative school placement, subsequent juvenile detention, and the salience of race. *Journal of Applied Research on Children: Informing Policy for Children at Risk, 5*(2), 14.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2007). *Woodcock-Johnson III tests of achievement* (3rd. ed.). Rolling Meadows, IL: Riverside.

Yoon, I. H. (2012). The paradoxical nature of whiteness-at-work in the daily life of schools and teacher communities. *Race, Ethnicity and Education, 15*(5), 587-613. doi: 10.1080/13613324.2011.624506

Younkin, P., & Kuppuswamy, V. (2017). The colorblind crowd? Founder race and performance in crowdfunding. *Management Science, 64*(7), 3269-3287. doi:10.1287/mnsc.2017.2774

Appendix A

Participant Invitation Letter

Good afternoon!

Are you concerned about the topic of *equity in school discipline*? If so, please consider participating in this survey research project. Your district and school administrator approved this research if you choose to participate.

The purpose of this study is to better understand if volunteer in-service teacher perception of the topic (equity in school discipline) changes over time based on different conditions. The study will take place over a 3-week period. Only in-service classroom teachers can participate. Each week will involve one activity. Each activity will include a small token of appreciation for providing your expert opinion/feedback/thoughts.

Participants will be offered the following 3 small tokens of appreciation for participating. For time period 1 = a jean pass. Time period 2 = an early leave pass. Time period 3 = a small blizzard (Dairy Queen) coupon.

Survey responses will be anonymous. No one, including the researcher, will be able to associate your responses with your identity. If you would like to participate, please complete the Microsoft form below. You will receive an email and individual link from Qualtrics next week. Thank you for taking the time to read this invitation.

Sincerely,

Nick Chastain

Doctoral Candidate, Valdosta State University

nrchastain@valdosta.edu

Microsoft Form Link (example)

If you received this email in error, please disregard.

Appendix B

Initial Interview Questions

Chastain, Nicholas interview questions:

Thank you for agreeing to participate in this portion of the study. This study was designed to look at an old problem in a new way. Specifically, I wanted to target in-service classroom teachers who were concerned about equity in school discipline... or, discipline disproportionality.

1. Did you complete the IAT assessment during the study? (experimental/control)
2. What are your thoughts on how the issues of race, culture, and bias impact your students (Blitz, Anderson, & Saastamoinen, 2016)?
3. What stands out to you most when reflecting on this experience?
4. [Experimental] What did you think about the IAT and the score you received?
   [Control] What did you think about the description of the IAT at time point 1? How do you think taking the IAT would have altered your experience?
5. [Experimental] Many participants did not complete the IAT during time point 2. Even fewer participants took the IAT during time point 3 but did complete the survey portion of the study... Why do you think participants completed the study without taking the IAT? Should this phenomenon guide future iterations of the study with other participants in different settings?
   [Control] Significantly more participants in the IAT group dropped out of the study than participants who did not take the IAT. How should this trend effect future iterations of this study with other participants in different settings?
6. [Experimental] Do you think the IAT was accurate? Some participants did not think the IAT was valid or reliable... do you agree? Why or why not? How did your opinion of the IAT change over time?
   [Control] Do you think implicit bias is measurable? Do you think implicit bias is real? How would you feel if given an IAT score indicating you held more or less implicit racial bias towards White or Black individuals than you believe?
7. [For both groups] What intervention strategy stood out most to you and why?
8. [For both groups] Have you been able to use any of the intervention strategies since participating in the study? If so, what did you think?
9. [For both groups] What was the driving force leading you to volunteer for this study? Were you familiar with the concept of implicit bias prior to participating in this study? If so, how did your perception of implicit bias change over time? If not, what are your thoughts on the topic after participating?
10. How do you see this experience changing your professional practice? How will this experience impact your daily life?
11. Many participants indicated training in implicit bias awareness and prevention would benefit their colleagues. How would training in this area effect the teaching and learning at your school? How will it effect teaching and learning in your classroom?
12. How does implicit bias effect the outcomes for all students in your schools? How does implicit bias effect the climate, culture, and number of office discipline referrals in your school? Looking back, do you think implicit bias ever effected any of these issues in your classroom? If so, how?

13. Although the findings of the quantitative portion of this study were non-significant, the feedback from many participants indicated this experience did affect them. How did this study alter your perception of equity in school discipline? How might this content be altered or improved upon in other settings and locations in the pre-K – 12 field?
14. Do you have any questions for me? Is there any topic you would like to discuss further? Will this experience change your professional practice? Thank you for your time and thoughts.

Appendix C

Adapted from Pepis (2017)

Unstructured interview Questions

1. I'm conducting these interviews so that I can learn from you about how it felt to participate in the online intervention. What stands out most in your mind when you think back on the experience?

2. If you received a score on the IAT, how did you feel about the score?

3. I don't want to know your IAT score unless you decide to share it with me. But I'm wondering if you remember what your score was?

4. Do you believe it was accurate? Why or why not?

5. Did your feelings about the IAT change after the intervention?

6. What do you think about the bias reduction strategies?

7. Were there any strategies in particular that stood out?

a. Why?

8. Have you had a chance to practice any of the strategies?

a. If so, elaborate on the experience

9. Had you heard of implicit bias before the intervention?

a. If so, what had you heard?

10. Do you think knowing about implicit bias will impact your teaching?

a. If so, how?

b. If not, why?

11. Did your views of how bias might affect your teaching change after taking the intervention?

12. Do you think there is implicit bias in schools? Can you think of any examples you might have seen during your field experience?

13. Additional questions will build on the participants' responses.

Appendix D

Preliminary Teacher Interview Questions and Presentation

You are being asked to participate in an interview as part of a research study entitled *"**Implicit Bias Awareness and Intervention Influence on In-service Classroom Teachers Promoting Equity in School Discipline: A Mixed Methods Study**,"* which is being conducted by Nicholas Chastain, a student at Valdosta State University. The purpose of the study is to better understand if making concerned in-service teachers aware of implicit bias and reviewing strategies designed to reduce bias effects their professional practice. You will receive no direct benefits from participating in this research study. However, your responses may help us learn more about possible improvements in how this content is delivered to other teachers in other places or settings. There are no foreseeable risks involved in participating in this study other than those encountered in day-to-day life. Participation should take approximately 30 minutes to 1 hour**.** The interviews will be recorded in order to accurately capture your concerns, opinions, and ideas. Once the recordings have been transcribed, the recordings will be destroyed. No one, including the researcher, will be able to associate your responses with your identity. Your participation is voluntary. You may choose not to participate, to stop responding at any time, or to skip any questions that you do not want to answer. You must be at least 18 years of age to participate in this study. Your participation in the interview will serve as your voluntary agreement to participate in this research project and your certification that you are 18 years of age or older.

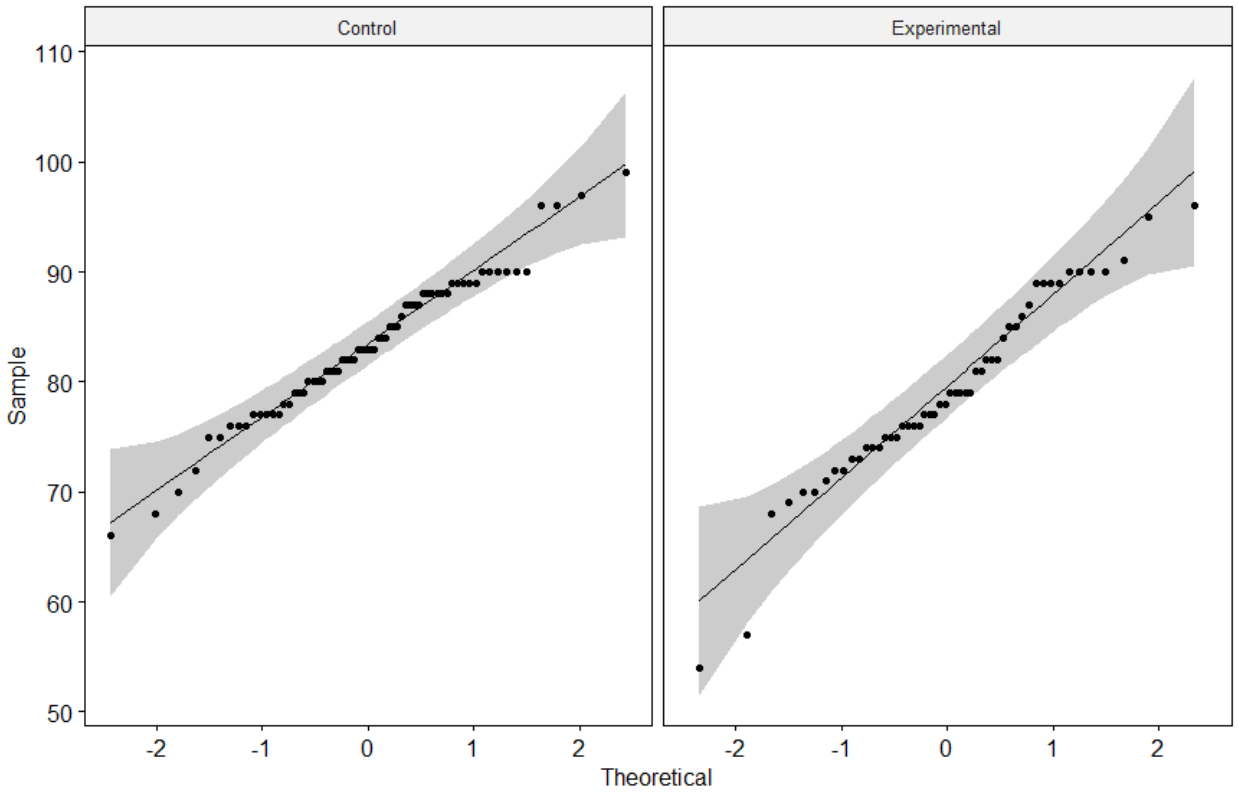
Questions regarding the purpose or procedures of the research should be directed to Nicholas Chastain at nrchastain@valdosta.edu. This study has been exempted from Institutional Review Board (IRB) review in accordance with Federal regulations. The IRB, a university committee established by Federal law, is responsible for protecting the rights and welfare of research participants. If you have concerns or questions about your rights as a research participant, you may contact the IRB Administrator at 229-253-2947 or irb@valdosta.edu.

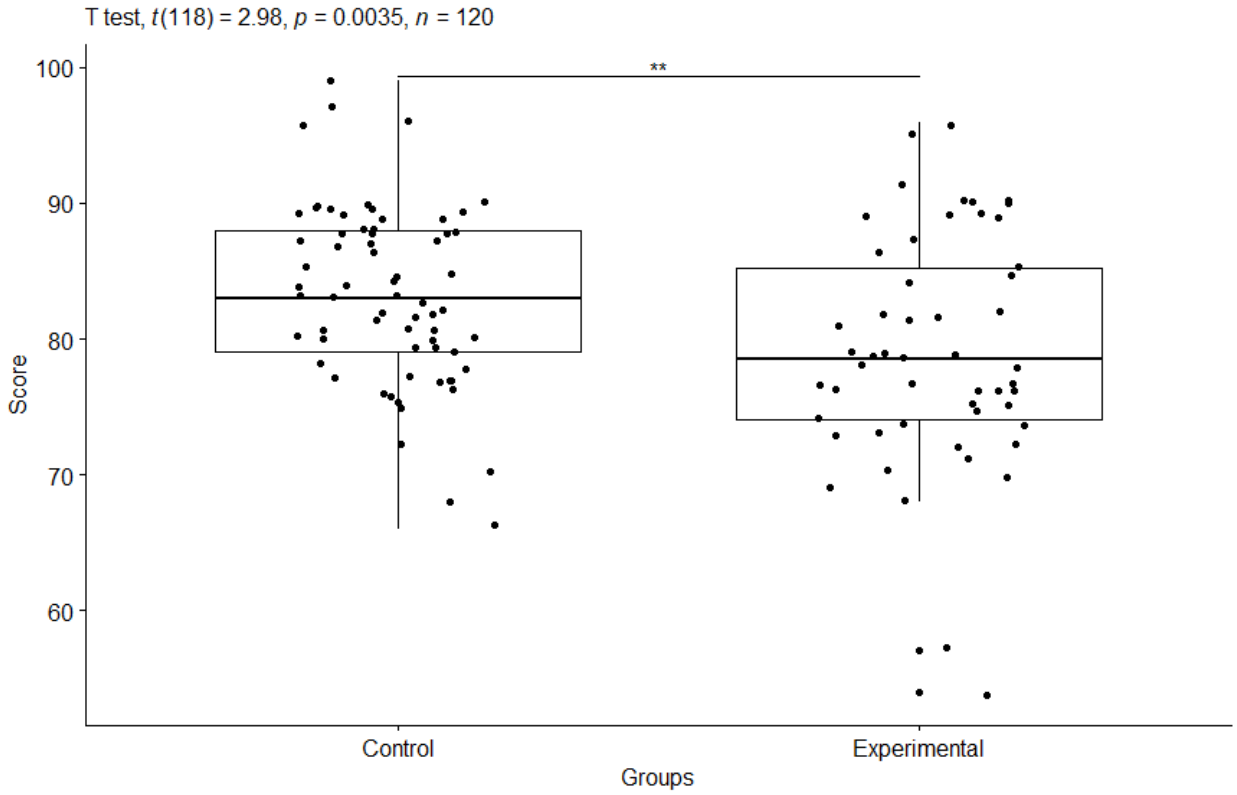
1. What stands out most in your mind when you think back on the experience?

2. I'm going to share some graphic output from the results of the quantitative portion of this study. If you have any questions, please don't hesitate to stop me.

The Teacher Multicultural Awareness Survey (TMAS) was administered to all teachers in this study at two time points. The survey was designed specifically for teachers. The 20 item instrument was scored on a 100-point scale. The higher the score, the more concern a teacher is thought to have about multicultural issues in their classroom. Do you have any questions?

3. There were two groups in this study. The control group did not take the Implicit Association Test (IAT). The experimental group did. Here is the graphic output of TMAS scores by group:





The results of the student *t* test indicated there was a significant difference between the control ($M = 83.2$, $SD = 6.59$) and experimental group ($M = 79.1$, $SD = 8.54$) ($t(118) = 2.98$, $p = .0035$; $d = 0.55$).

3a. Are you surprised by this result? If so, why? If not, why?

The control group pretest and posttest scores were analyzed with a dependent means *t* test. There was no significant difference in control group scores:

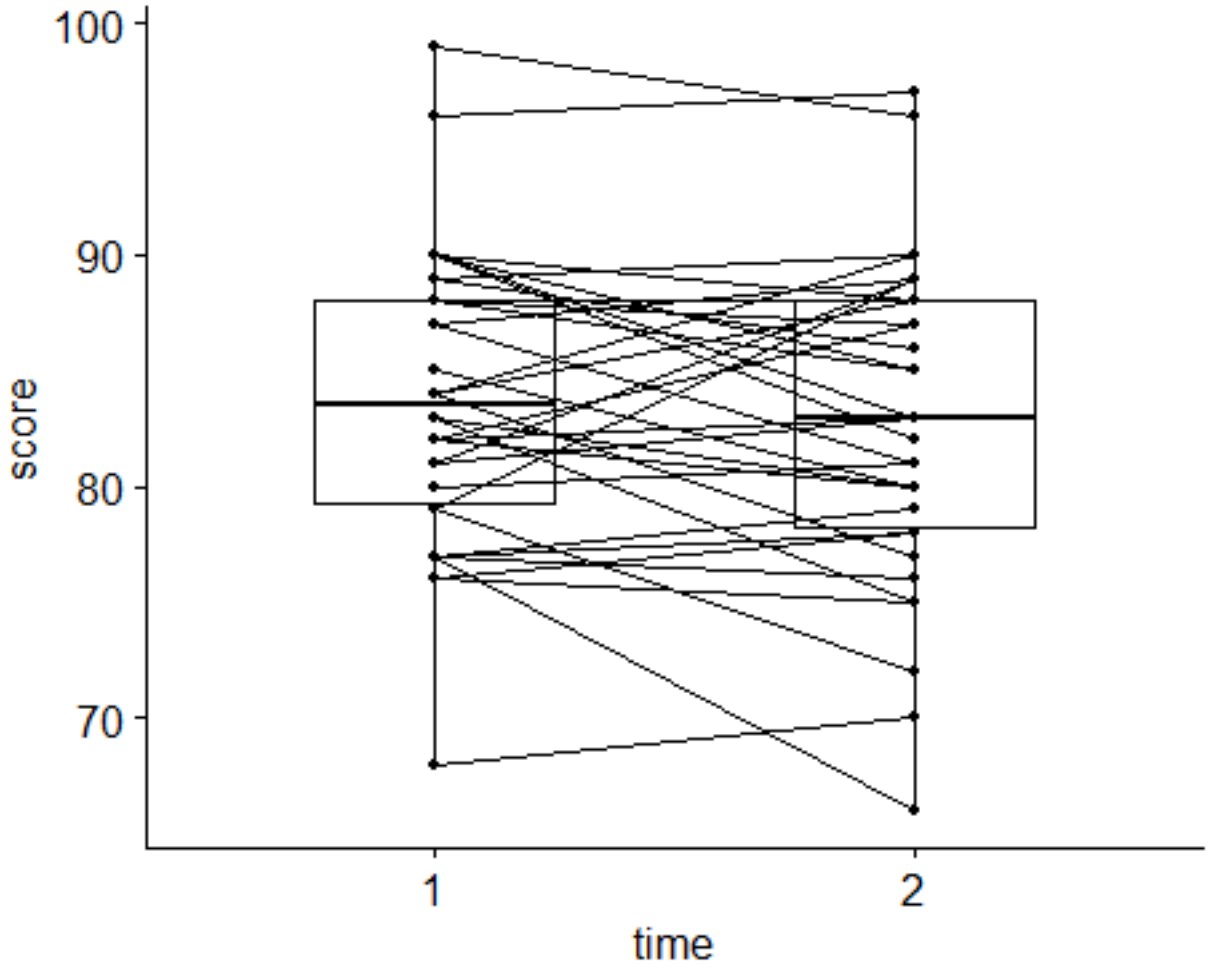


Figure 1 Control Group pretest and posttest scores

The results of the dependent *t* test indicated there was not a significant difference between the pretest ($M = 83.8$, $SD = 6.29$) and posttest ($M = 82.7$, $SD = 6.93$) ($t(33) = -1.25$, $p = .22$; $d = -.22$) control group TMAS scores. Cohen's *d* (effect size) indicates a small negative affect on control group TMAS scores.

3b. Are you surprised about these results?

There was no significant difference in experimental group scores, either:

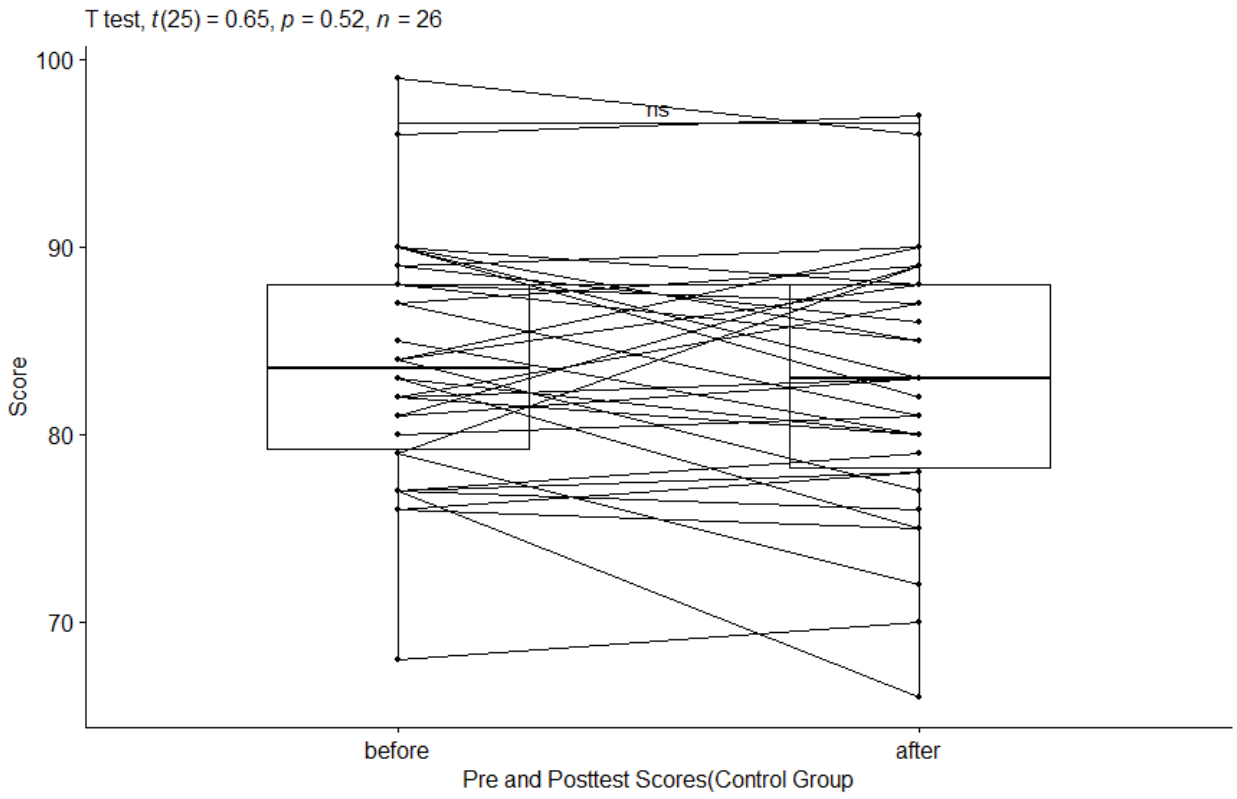
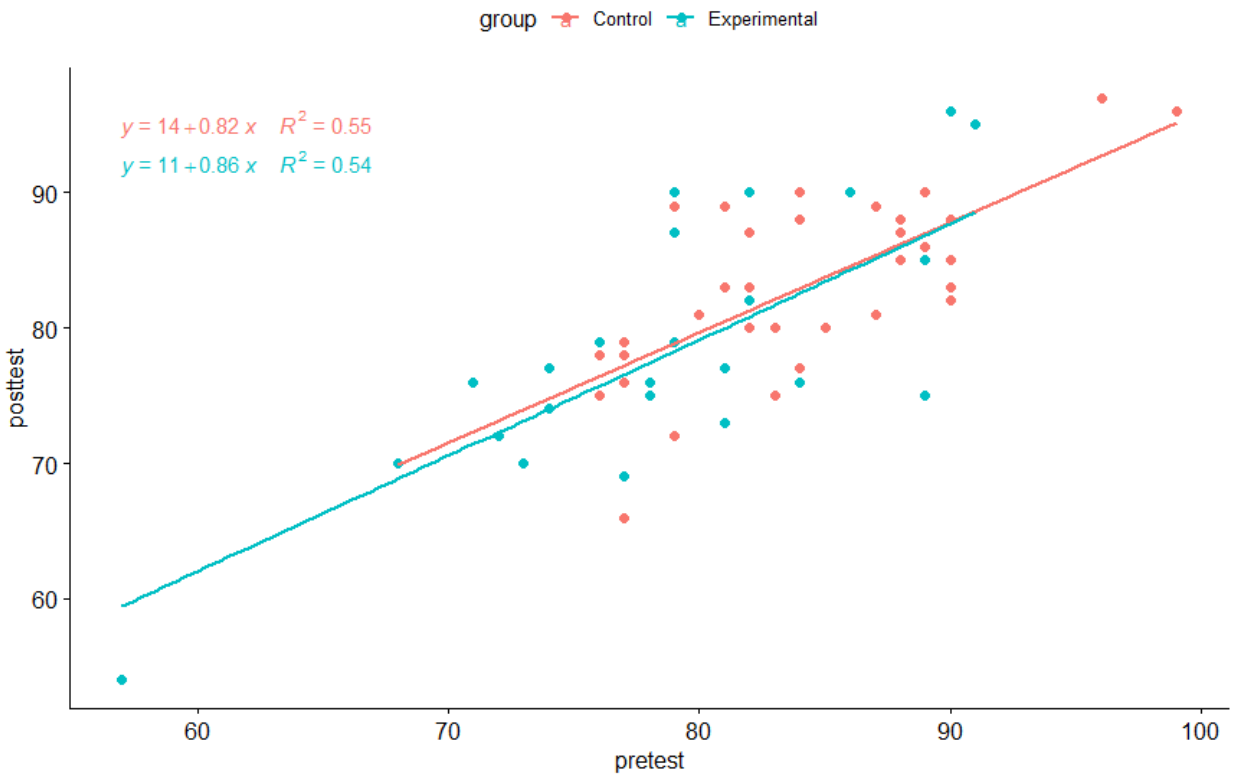


Figure 2 Experimental Group pretest and posttest scores

The results of the dependent *t* test indicated there was not a significant difference between the pretest ($M = 79.5$, $SD = 7.94$) and posttest ($M = 78.7$, $SD = 9.23$) ($t(25) = 0.65$, $p = 0.52$; $d = .13$) experimental group TMAS scores. Cohen's *d* (effect size) indicates a negligible positive effect for experimental group TMAS scores.

3b. Are you surprised about these results?

Further analysis included an ANCOVA with the pretest score as a covariate:



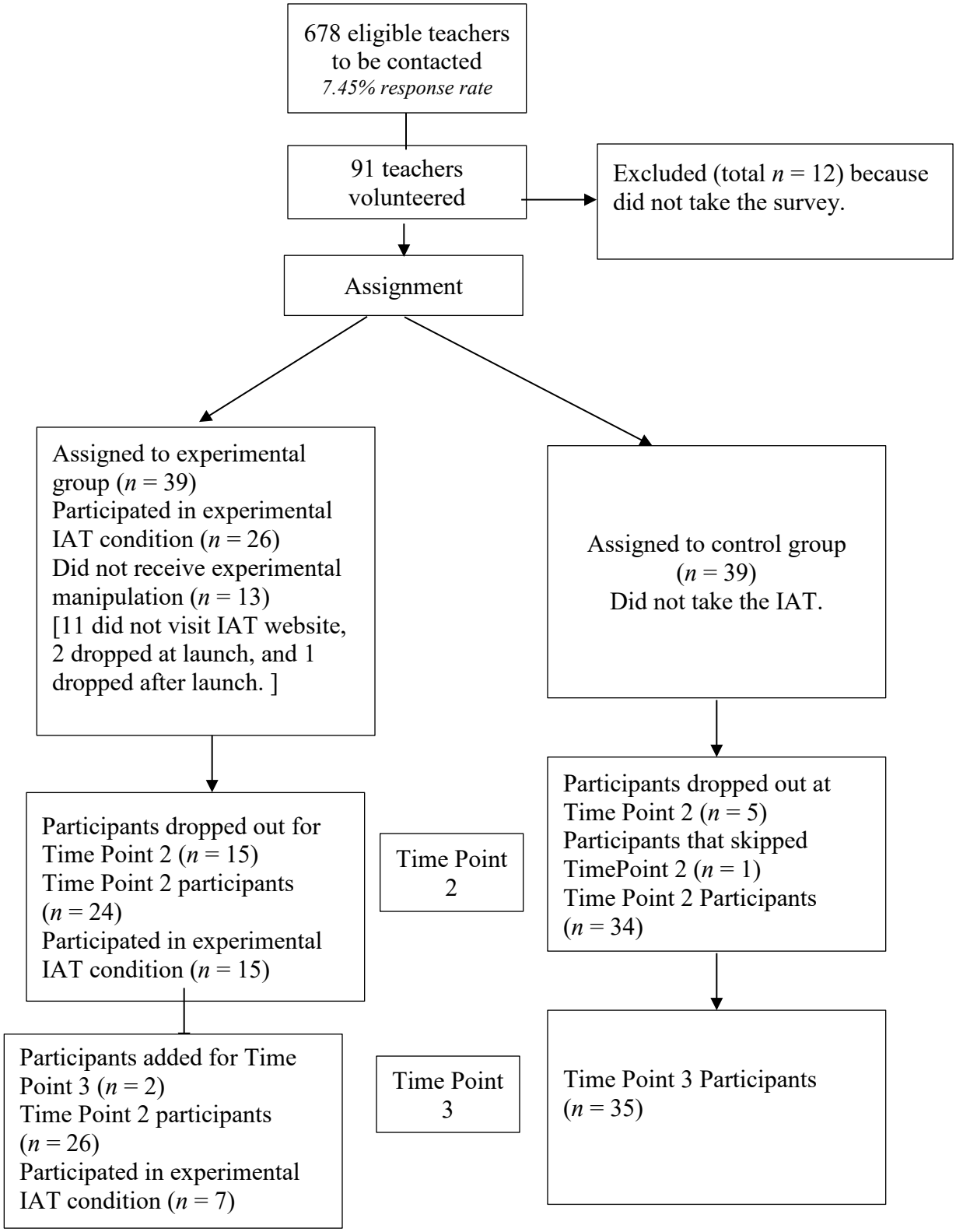
After adjustment for pretest TMAS score, there was not a statistically significant difference in posttest TMAS score between the groups, $F(1, 57) = 0.10$, $p > 0.05$.

3c. Do you have any questions about the graphics or results so far?

4. Did you take the IAT? Experimental group participants took the IAT 3 times during the study. They were randomly assigned to the experimental group by Qualtrics software. Here is the breakdown of participants flowing through the study:

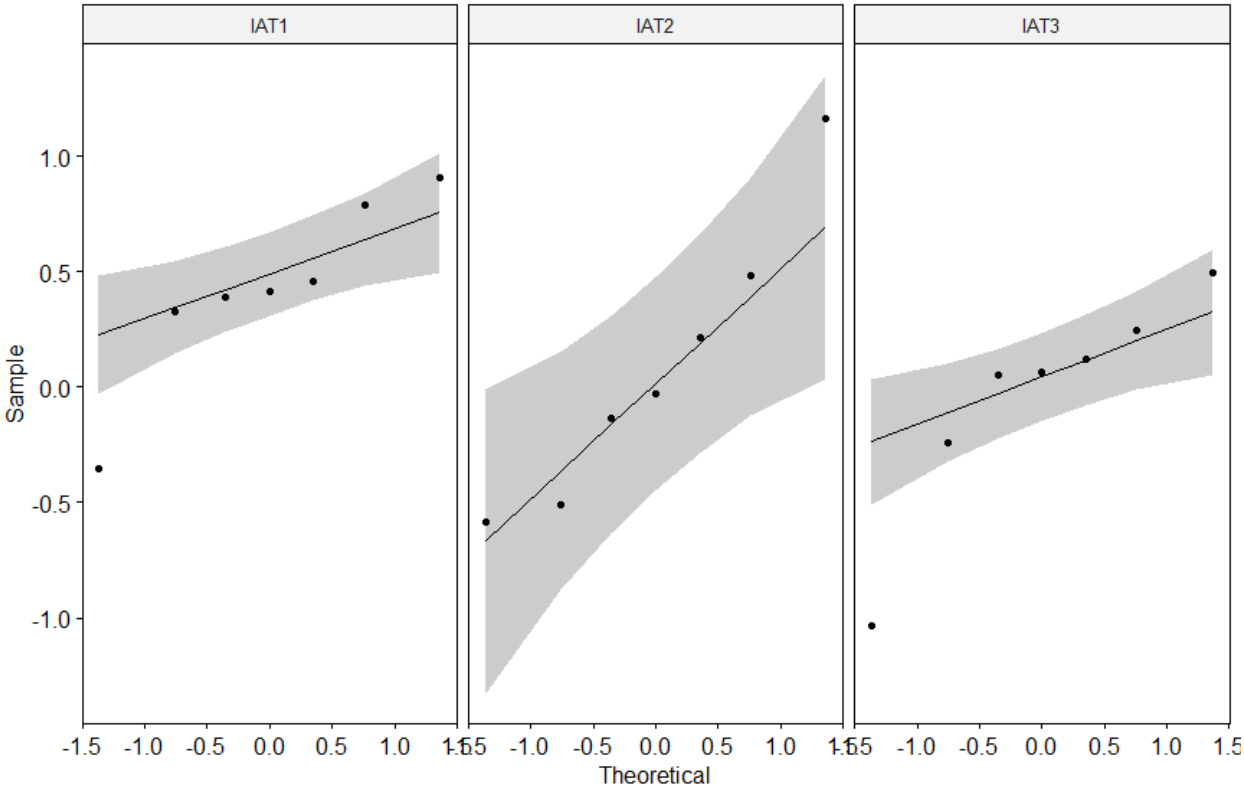
Figure 1

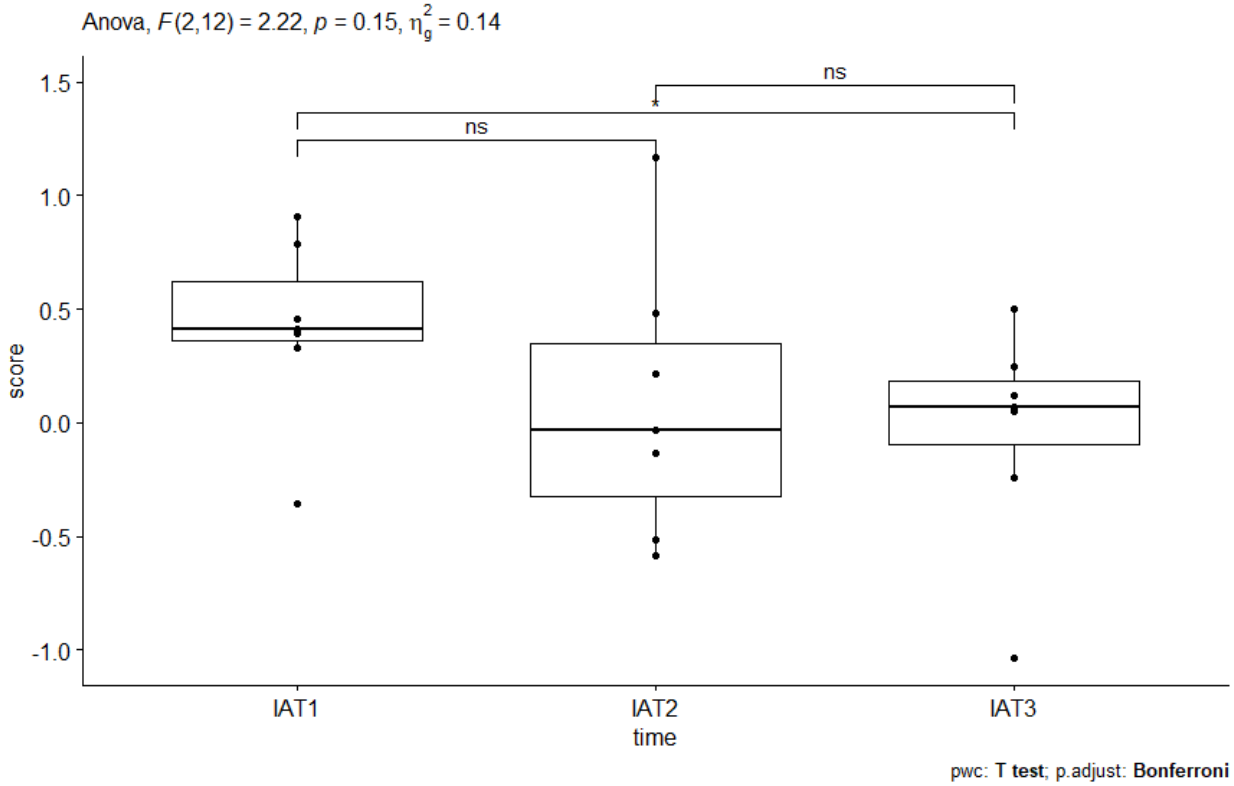
Flow of participants for the given study procedures, assignment (time point 1), time point 2, and time point 3.



4a. Do you have any thoughts, comments or questions on the participant flow chart (differences in group retention etc.)?

Of the participants who completed all three IAT assessments, this is what their scores looked like over time. A repeated measures ANOVA calculated if there was a significant difference in participant scores over time.

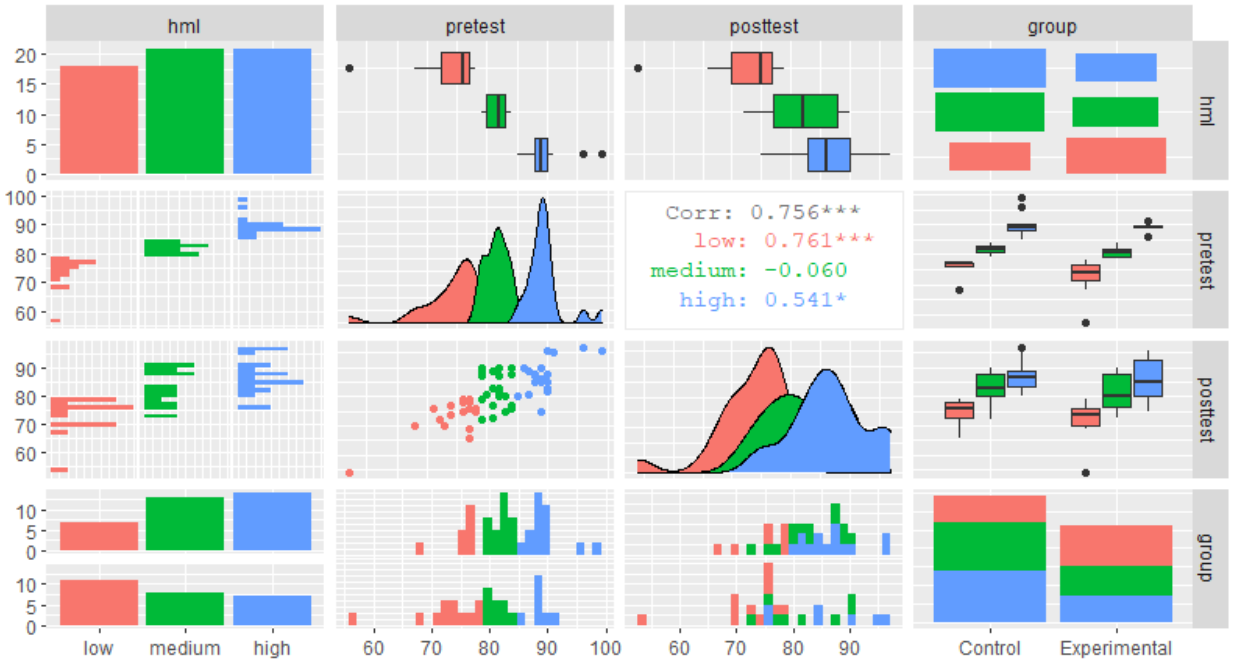


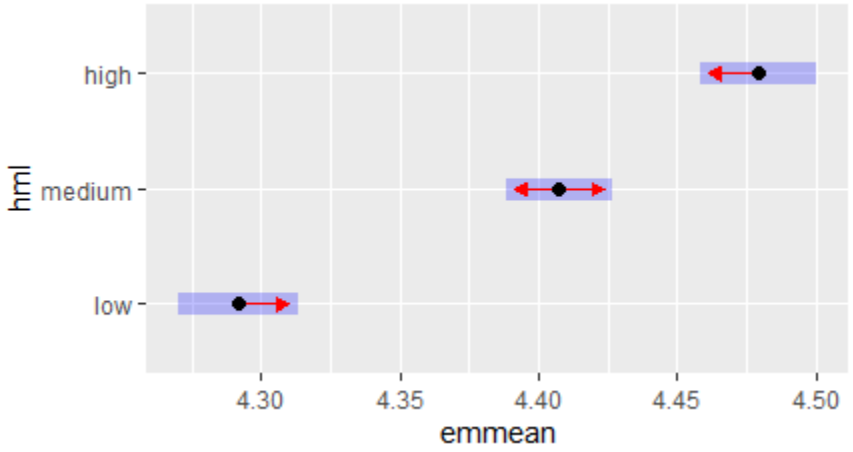


The scores did trend down over time. The results indicated participant IAT scores were significantly different between time point 1 and 3 ($t(6) = 4.07$, $p = .007$), but the overall results were non-significant ($F(2, 22) = 5.43$, $p = 0.15$) with a small positive effect size ($\eta^2 = .04$).

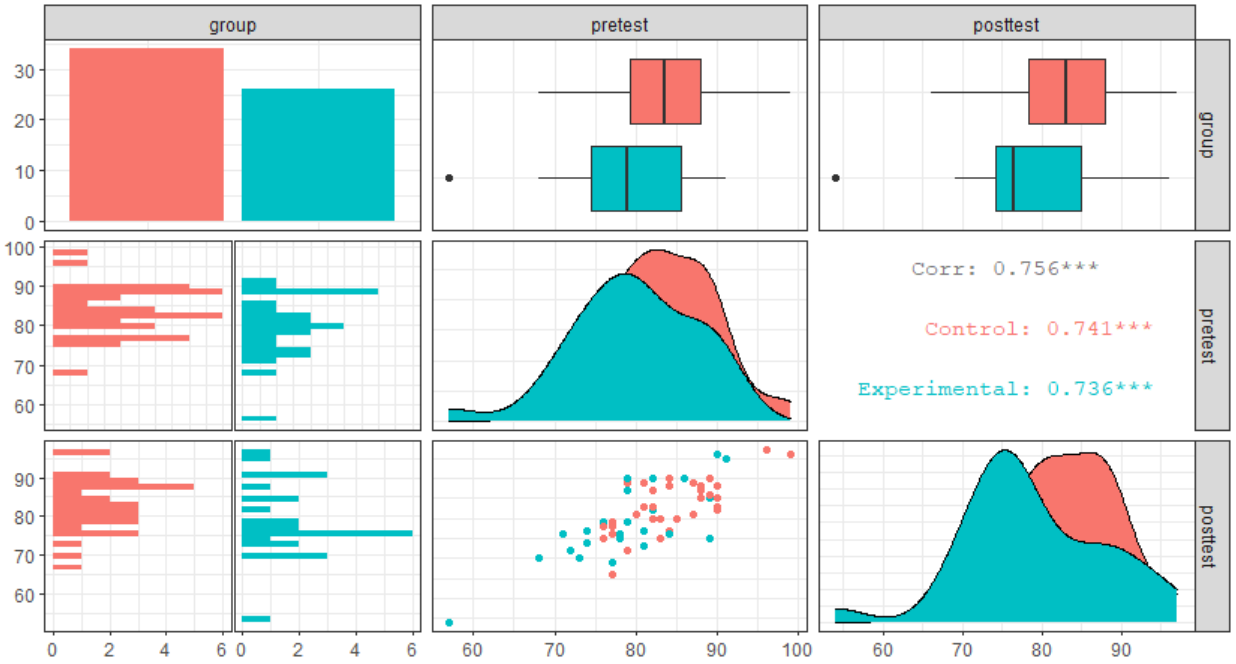
4b. Do you have any thoughts on the IAT scores? Are the results surprising? Why?

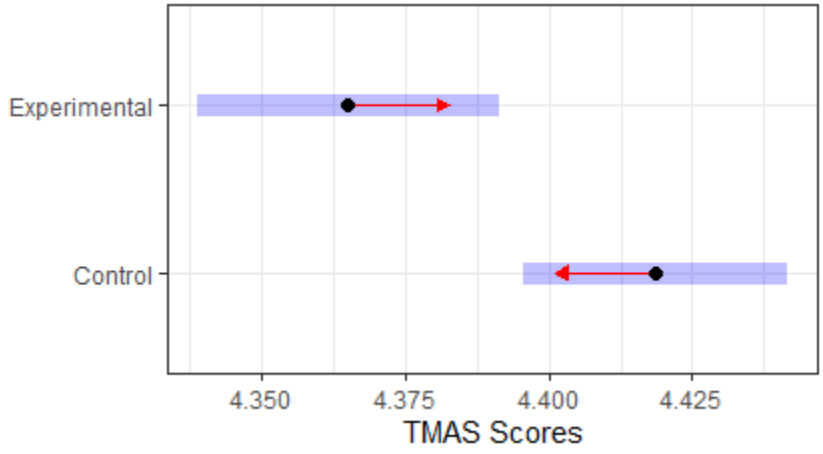
Here is the graphic output of participants who scored high, medium, or low on the pretest TMAS: red group scored low, green medium, and blue high on the pretest TMAS.





And here is the correlation table of pretest and posttest scores: The red group was control, aqua experimental.





5. The findings in the quantitative portion of this study were largely non-significant (statistically). The feedback from many participants indicated this experience was personally significant. How did this study alter your views on equity in school discipline?

6. What are your thoughts on how the issues of race, culture, and bias impact your students (Blitz, Anderson, & Saastamoinen, 2016)?

7. What stands out to you most about the quantitative results of this study?

8. Experimental Group: What was taking the IAT like?

Control Group: What do you think taking the IAT would have been like?

10. What were the driving forces that led to you to participate in this study?

11. Many participants shared detailed stories about personal experiences with implicit bias. Have you encountered any situations or experiences since participating you would like to describe?

13. Would training in this area effect the teaching and learning at your school? How will it affect teaching and learning in your classroom?

14. Do you have any questions about this topic or the research presented? Is there any topic or question you would like to revisit or answer you would like to clarify?

Thank you for participating in this study. If you have any questions, please email nrchastain@valdosta.edu.

Appendix E

Final Interview Questions Based on Participant Feedback

You are being asked to participate in an interview as part of a research study entitled *"**Implicit Bias Awareness and Intervention Influence on In-service Classroom Teachers Concerned about Equity in School Discipline: A Mixed Methods Study**,"* which is being conducted by Nicholas Chastain, a student at Valdosta State University. The purpose of the study is to better understand if making concerned in-service teachers aware of implicit bias and reviewing strategies designed to reduce bias effects their professional practice. You will receive no direct benefits from participating in this research study. However, your responses may help us learn more about possible improvements in how this content is delivered to other teachers in other places or settings. There are no foreseeable risks involved in participating in this study other than those encountered in day-to-day life. Participation should take approximately 30 minutes to 1 hour**.** The interviews will be recorded to accurately capture your concerns, opinions, and ideas. Once the recordings have been transcribed, the recordings will be destroyed. No one, including the researcher, will be able to associate your responses with your identity. Your participation is voluntary. You may choose not to participate, to stop responding at any time, or to skip any questions that you do not want to answer. You must be at least 18 years of age to participate in this study. Your participation in the interview will serve as your voluntary agreement to participate in this research project and your certification that you are 18 years of age or older.

Questions regarding the purpose or procedures of the research should be directed to Nicholas Chastain at nrchastain@valdosta.edu. This study has been exempted from Institutional Review Board (IRB) review in accordance with Federal regulations. The IRB, a university committee established by Federal law, is responsible for protecting the rights and welfare of research participants. If you have concerns or questions about your rights as a research participant, you may contact the IRB Administrator at 229-253-2947 or irb@valdosta.edu.

Definition of key terms:

Teacher Multicultural Awareness Survey (TMAS) – 20 instrument Likert type questionnaire designed specifically for classroom teachers. Scores range from 20-100. 20 would indicate no concern or belief that multicultural issues in the classroom are important. 100 would mean highly concerned.

Race Implicit Association Test (IAT): A latency based cognitive assessment designed to measure an individual implicit racial biases based on white and black skin tone, respectively. A score of +2 would mean strong preference for white skin. -2 would mean a strong preference for black skin. 0.0 would mean no preference for white or black skin.

Timeline of Quantitative portion of the study:

Week 1: Demographic information, TMAS, Devine intervention for control group

Week 1: Demographic information, TMAS, IAT, Devine intervention for experimental group

Week 2: COVID questionnaire for control group

Week 2: COVID questionnaire and IAT for experimental group

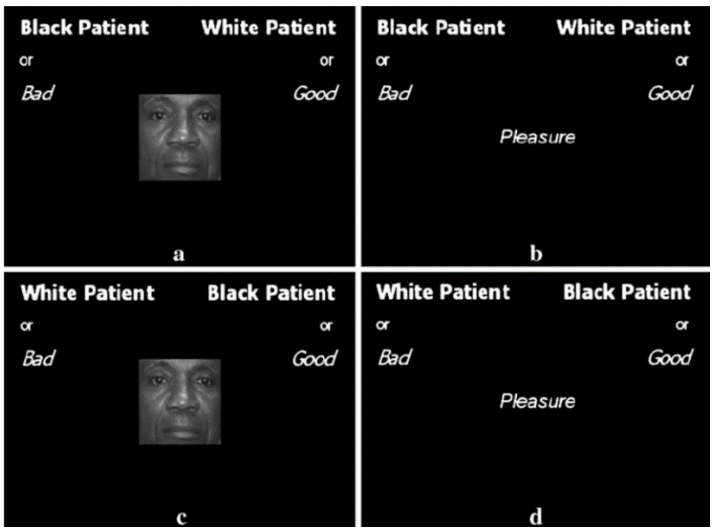
Week 3: TMAS and follow up demographic questions for control group

Week 3: IAT, TMAS, and follow up demographic questions for experimental group

All volunteers were emailed a secure link to a survey via Qualtrics. Section 1 obtained demographic information. Section 2 all participants took a 20-item survey: The Teacher Multicultural Awareness Survey (TMAS). The TMAS was specifically designed for classroom teachers and produces scores ranging from 20 – 100. Higher scores indicate teachers have more awareness and appreciation of multicultural issues in the classroom.

Lower scores the opposite. Most classroom teachers in the United States score around 80 – 81 [grand mean] (out of 100). The mean score of all participant TMAS assessments in this study was 81.5. This includes pretest AND posttest scores.

Once the pretest TMAS was complete, participants were randomly assigned to either the experimental or control group. The experimental group was asked to complete the race Implicit Association Test (IAT). Did you take the IAT? If not, let me explain... The IAT is a latency-based instrument. Individuals are asked to hit a computer key as fast as possible when they see words that are good or bad, or faces that are black or white. It looks similar to this picture:



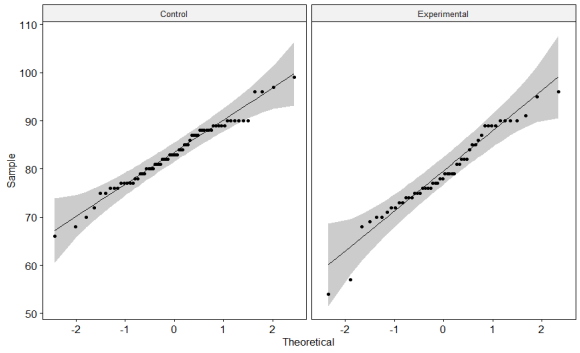
If an individual takes too long to hit a computer key, their scores are invalidated. The cognitive latency based assessment measures bias by measuring the latency in hitting the computer keys. On the race IAT, positive scores mean preference for white skin, negative scores a preference for black skin.

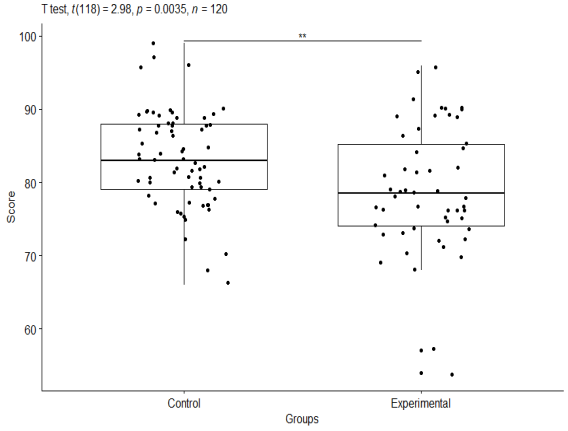
Time point 2: Participants from both groups were asked to login one week later. Control group participants answered a question about the COVID-19 pandemic. Experimental group participants answered the same question but were then asked to complete another race IAT.

Time point 3: Participants were asked to login in one last time. Control group participants took the posttest TMAS and follow up demographic questions. Experimental group participants took the IAT, the posttest TMAS and follow up demographics.

1. What stands out most in your mind when you think about participating?

TMAS pretest and posttest scores by control group and experimental group:





The results of the student *t* test indicated there was a significant difference between the control (M = 83.2, SD = 6.59) and experimental group (M = 79.1, SD = 8.54) ($t(118) = 2.98$, $p = .0035$; $d = 0.55$). The control group mean score was significantly higher than experimental group mean score on the TMAS.

The independent t-test compared all the TMAS scores (pre and posttest) of the control group and experimental group. The control group scored significantly higher than the experimental group. This indicates the control group had more concern over multicultural issues than the experimental group.

2. What are your thoughts on why the mean TMAS scores were different between the control and experimental groups? Do the results surprise you?

The results of the dependent *t* test indicated there was not a significant difference between the pretest (M = 83.8, SD = 6.29) and posttest (M = 82.7, SD = 6.93) ($t(33) = -1.25$, $p = .22$; $d = -.22$) control group TMAS scores. Cohen's d (effect size) indicates a small negative affect on control group TMAS scores. The control groups mean score went down by 1.1 point from pretest to posttest. Participants scores indicate they had less appreciation and awareness of multicultural issues in the classroom by participating in this study.

There was no significant difference in experimental group scores, either:

The results of the dependent *t* test indicated there was not a significant difference between the pretest (M = 79.5, SD = 7.94) and posttest (M = 78.7, SD = 9.23) ($t(25) = 0.65$, $p = 0.52$; $d = .13$) experimental group TMAS scores. Cohen's d (effect size) indicates a negligible positive effect for experimental group TMAS scores. The experimental group mean score did not decrease as much (-0.8) by participating in the study.

The control group pre and post test scores were not significantly different.

The experimental group pre and posttest scores were not significantly different.

The results of a one-way ANCOVA reduced the effect of differences between the groups on the pretest TMAS. This analysis did not find a significant difference between the control group and experimental group on TMAS scores.

3. Do you have any questions about the results so far?

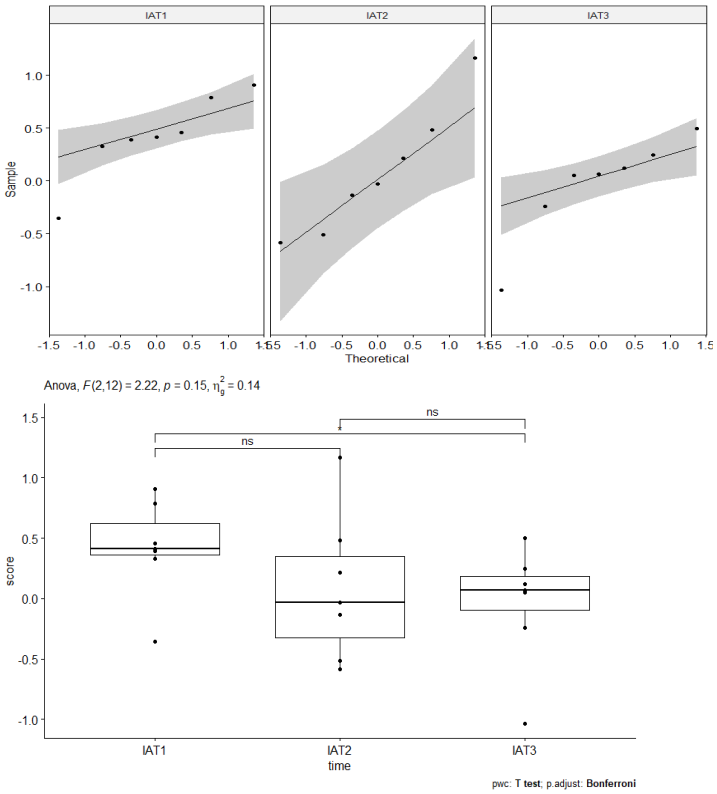
The number of participants in each group was equal at time point 1 (39 in each).

Only 5 participants dropped out of the study from the control group (34 completed the study).

13 participants dropped out of the experimental group completely. 32 participants did not complete the IAT portion of the study (26 partially completed, 7 fully completed the study).

4. Does that dropout rate difference seem reasonable/explainable to you?

The IAT scores for those 7 participants did trend down over time (time point 1 mean = 0.42, 2 mean = 0.09, and 3 mean = -0.04). This indicates participants went from a moderate preference for white faced people at time point 1 to a negligible preference for black faced people at time point 3 as measured by the race IAT...



The results indicated participant IAT scores were significantly different between time point 1 and 3 ($t(6) = 4.07$, $p = .007$), but the overall results were non-significant ($F(2, 22) = 5.43$, $p = 0.15$) with a small positive effect size ($\eta^2 = .04$).

5. What are your thoughts on this change in preference?

6. A number of participants provided feedback this experience was *personally* significant. How did this study alter or shape your feelings on the topic of "equity in school discipline"? Do you think the topic is more or less important for k-12 students and teachers after participating?

7. How do the issues of race, culture, and bias impact your students (Blitz, Anderson, & Saastamoinen, 2016)?

8. What was the driving force that led to you to participate in this study?

9. Are there any stories or personal experiences with implicit bias in the k-12 setting you can recall? Are you willing to describe an experience or event at a school you worked at related to implicit bias and equity in school discipline? How did this experience affect you? How did it affect the student?

10. How should K-12 leaders address the issue of equity in school discipline with their teachers?

11. Would training in this area effect the teaching and learning at your school? Teaching and learning in your classroom?

12. Do you have any questions about this topic, or the research presented? Is there any topic or question you would like to revisit or answer you would like to clarify?

Thank you for participating in this study. If you have any questions, please email nrchastain@valdosta.edu.

Appendix F

Descriptive Statistics for all TMAS Scores at Baseline

Table 6

Descriptive Statistics of Baseline Control and Experimental Group Mean Pretest TMAS Scores

Group	Time	n	*M*	*SD*	Min[a]	Max[b]	Skewness	Kurtosis
Both	Pretest	78	81.19	8.1	57	99	-0.53	0.27
Control	Pretest	39	82.87	7.5	60	99	-0.67	1.02
Experimental	Pretest	39	79.51	8.43	57	96	-0.34	-0.25

Note. [a]Minimum value; [b]Maximum value.

Appendix G

Institutional Review Board Protocol Exemption Report

Institutional Review Board (IRB)
For the Protection of Human Research Participants

PROTOCOL EXEMPTION REPORT

Protocol Number: 04017-2020

Responsible Researcher: Nicholas Chastain

Supervising Faculty: Dr. James Pate

Project Title: *Equity in School Discipline.*

INSTITUTIONAL REVIEW BOARD DETERMINATION:

This research protocol is **Exempt** from Institutional Review Board (IRB) oversight under Exemption **Category 2**. Your research study may begin immediately. If the nature of the research project changes such that exemption criteria may no longer apply, please consult with the IRB Administrator (irb@valdosta.edu) before continuing your research.

ADDITIONAL COMMENTS:

- *Upon completion of this research study all data (email correspondence, survey data, participant lists, etc.) must be securely maintained (locked file cabinet, password protected computer, etc.) and accessible only by the researcher for a minimum of 3 years.*

☒ ***If this box is checked, please submit any documents you revise to the IRB Administrator at irb@valdosta.edu to ensure an updated record of your exemption.***

Elizabeth Ann Olphie *04.08.2020*
Elizabeth Ann Olphie, IRB Administrator

Thank you for submitting an IRB application.
Please direct questions to irb@valdosta.edu or 229-253-2947.

Revised: 06.02.16

Appendix H

Expert Panel Review Form

Expert Panel Review

Dear Expert Reviewer,

I need your help! Your knowledge and expertise about this topic as a leader in the field of education will provide verification of items in the interview portion of the research study - "Implicit Bias Awareness and Intervention Influence on In-Service Classroom Teacher Promoting Equity in School Discipline: A Mixed Methods Study". Your feedback will help improve this portion of the study by rewording items, removing items, or including additional items. Your help is essential and I appreciate the time you are taking to examine the questions for me. You will receive an invitation that will open this secure link to view the questions and content: https://docs.google.com/document/d/1Cv4jIJJuMtzmBc_olPjLAYsTCiQY60KNW6OK7RelqV8/edit?usp=sharing

Please review the questions and content before proceeding.

Directions: Please select the option that best represents your response. If you answer "no" to items 1, 2, 4, 5, or 6 please supply an explanation in the space provided. However, if you answer "Yes" to items 3 or 7, please provide an explanation.

* Required

1. Do you think the items matched the stated purpose of this study? *

Yes
No

If no, please explain.

Your answer

2. Is the presentation of statistical information and graphical output clear? *

Yes
No

If no, please explain.

Your answer

3. Are any of the questions leading, culturally insensitive, or otherwise inappropriate? *

Yes
No

If yes, please explain.

Your answer

4. Is each item understandable? *

Yes
No

If no, please explain.

Your answer

5. Is each item unambiguous? *

Yes
No

If no, please explain.

Your answer

6. Is each item grammatically correct? *

Yes
No

If no, please explain.

Your answer

7. Is there any section you feel requires additional information or additional items to improve the presentation and questions? *

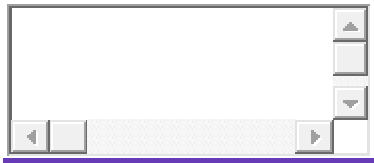
Yes
No

If Yes, please explain.

Your answer

Thank you for the time and effort! Please leave any personal thoughts about this study below. If you received this form - you are someone I have the utmost respect for. Thank you for helping me reach this point. -Nick

Your answer



Submit

Appendix I

Feedback from Expert Review Panelist

Dr. Blank (omitted for privacy),

In a nutshell, participants did not change in a statistically significant way on pretest, posttest, and repeated measure assessment scores on the TMAS (Teacher Multicultural Awareness Survey), or the IAT (Implicit Association Test). The written responses they provided within the survey platform, however, indicate many of them DID change. The graphs and stat's below are going to take about 60 pages of dissertation space to explain correctly. But they can paint a pretty good picture of how the "stats" looked. Trying to figure out the disconnect between assessment scores and what respondents felt, is where I am at. Thank you for your help. I really, really, appreciate it. –Nick

OK… so I have made some suggestions. Honest ones that I hope are helpful. It is hard to take just bits and pieces and totally understand what you are writing/researching, but I think I get the gist enough maybe to give a little advice.

When it came to my interview questions, I can tell you that I simplified, and then I simplified again. The wording became more and more simplistic, and this allowed for the voices of the people to come through. When my "academic" voice was written into the language of the question, even that guided people into how they tried to answer (I found as I did work before I ever started the actual dissertation process) and I discovered if I wanted to hear the person, their voice, then I had to try to be as generic as possible (basically shut up…on paper). Ask in simply stated questions that got to the heart of what I needed to know in order to pull the data into some sort of organizational pattern that I could use to draw some conclusions from and then also use their voice to substantiate it. Ask as simply as possible exactly what you need to know in order to elicit the answers that will substantiate the data you have. A HARD thing to do, but the questions are really important, because they do pull out of the person what helps to support the data…OR…they lead them down the wrong path or confuse/intimidate them and prevent them from giving you what you need to finish the task.

Does that make any sense?? I highlighted in green

My first draft of questions:

Chastain, Nicholas interview questions:

Thank you for agreeing to participate in this portion of the study. This study was designed to look at an old problem in a new way. Specifically, I wanted to target in-

service classroom teachers who were concerned about equity in school discipline... or, discipline disproportionality.

1. Did you complete the IAT assessment during the study? (experimental/control)
2. What are your thoughts on how the issues of race, culture, and bias impact your students (Blitz, Anderson, & Saastamoinen, 2016)?
3. What stands out to you most when reflecting on this experience?
4. [Experimental] What did you think about the IAT and the score you received? [Control] What did you think about the description of the IAT at time point 1?

   How do you think taking the IAT would have altered your experience?

5. [Experimental] Many participants did not complete the IAT during time point 2. Even fewer participants took the IAT during time point 3 but did complete the survey portion of the study... Why do you think participants completed the study without taking the IAT? Should this phenomenon guide future iterations of the study with other participants in different settings?
   [Control] Significantly more participants in the IAT group dropped out of the study than participants who did not take the IAT. How should this trend effect future iterations of this study with other participants in different settings? *I would omit this question. It is too in depth in thought for a single question (I think), and asks participants to evaluate the behaviors of others when you are really wanting to determine what causes your data not to match up for each individual. That needs to be the focus of every question—getting to the core of why that is happening for each participant. If you branch out to far beyond just that focus, it is going to make the paper more than what you want to handle. Without saying it like this, you basically have to go to your people and say—your test score(s) say one thing, but your written responses show another. Can you possible explain why? They know why… and you have to figure out a way in your questioning techniques to lead them to talk about that without hemming them in (or limiting them in how they respond) with the questions of your interview, which is what I*

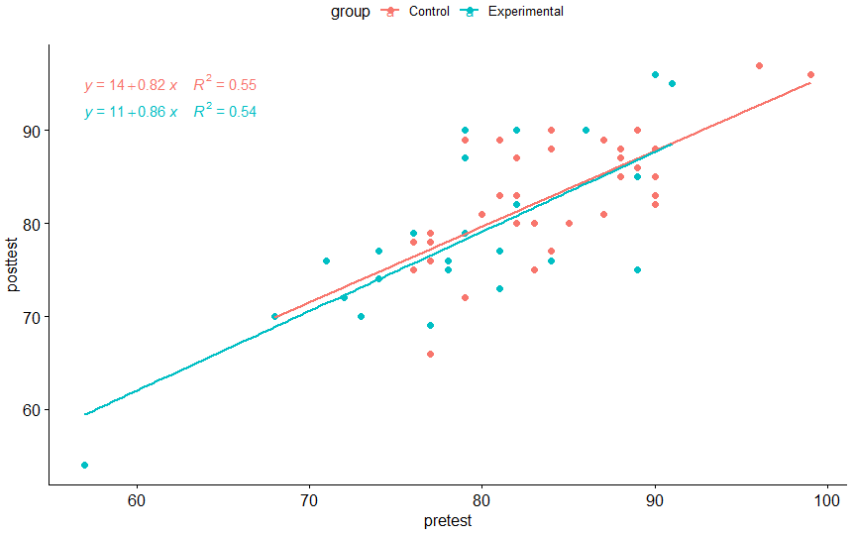
suspect the questions of the pre, post, etc. assessments did and thus the test scores did not vary.

6. [Experimental] Do you think the IAT was accurate? Some participants did not think the IAT was valid or reliable... do you agree? Why or why not? How did your opinion of the IAT change over time? *I might omit this sentence…If you want to see if the individual participant feels if the test is valid or reliable, but again… this one question now asks accurate/valid/reliable/measurable/is it real/… do not overload if you want people to open up. Try to word the question in fewer words and allow the interviewee to do the talking*
   [Control] Do you think implicit bias is measurable? Do you think implicit bias is real? How would you feel if given an IAT score indicating you held more or less implicit racial bias towards White or Black individuals than you believe?
7. [For both groups] What intervention strategy stood out most to you and why?
8. [For both groups] Have you been able to use any of the intervention strategies since participating in the study? If so, what did you think? *Describe and experience you had implementing one of the intervention strategies, if applicable….? Is the goal to be open ended so the interviewee shares more?*
9. [For both groups] What was the driving force leading you to volunteer for this study? Were you familiar with the concept of implicit bias prior to participating in this study? If so, how did your perception of implicit bias change over time? If not, what are your thoughts on the topic after participating?
10. How do you see this experience changing your professional practice? How will this experience impact your daily life? *Again, and maybe this is me, but wordy question? Clean them up and open doors for interviewees to talk*
11. Many participants indicated training in implicit bias awareness and prevention would benefit their colleagues. How would training in this area effect the teaching and learning at your school? How will it affect? teaching and learning in your classroom?
12. How does implicit bias effect the outcomes for all students in your schools? How does implicit bias effect the climate, culture, and number of office discipline referrals in your school? Looking back, do you think implicit bias ever effected any of these issues in your classroom? If so, how?
13. Although the findings of the quantitative portion of this study were non-significant, the feedback from many participants indicated this experience did affect them. How did this study alter your perception of equity in school discipline? How might this content be altered or improved upon in other settings and locations in the pre-K – 12 field?
14. Do you have any questions for me? Is there any topic you would like to discuss further? Will this experience change your professional practice? Thank you for your time and thoughts.

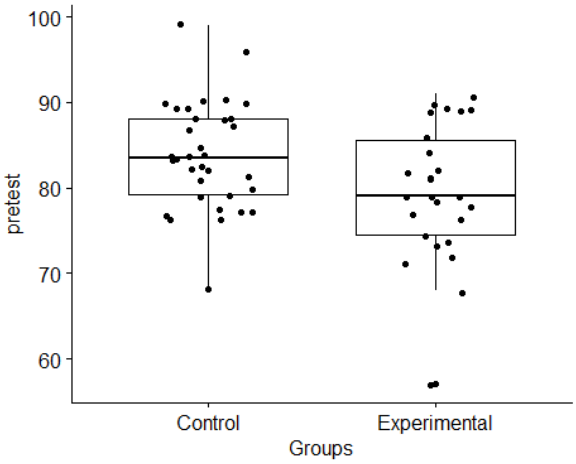
Question 13 is where the interview needs to go earlier. It is what you need to know in order to properly discuss the data…right? I think from what I understand. So, write questions that center on this. Turn this one into two. Some of the earlier ones can be re-written to focus more on this idea.

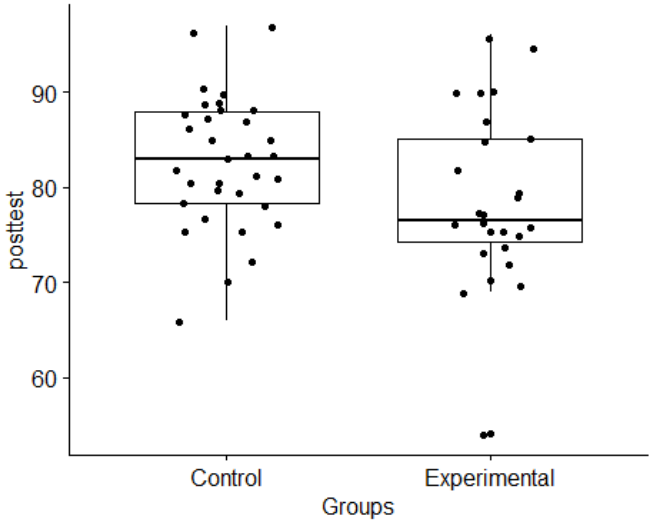
Descriptive statistics and graphs:

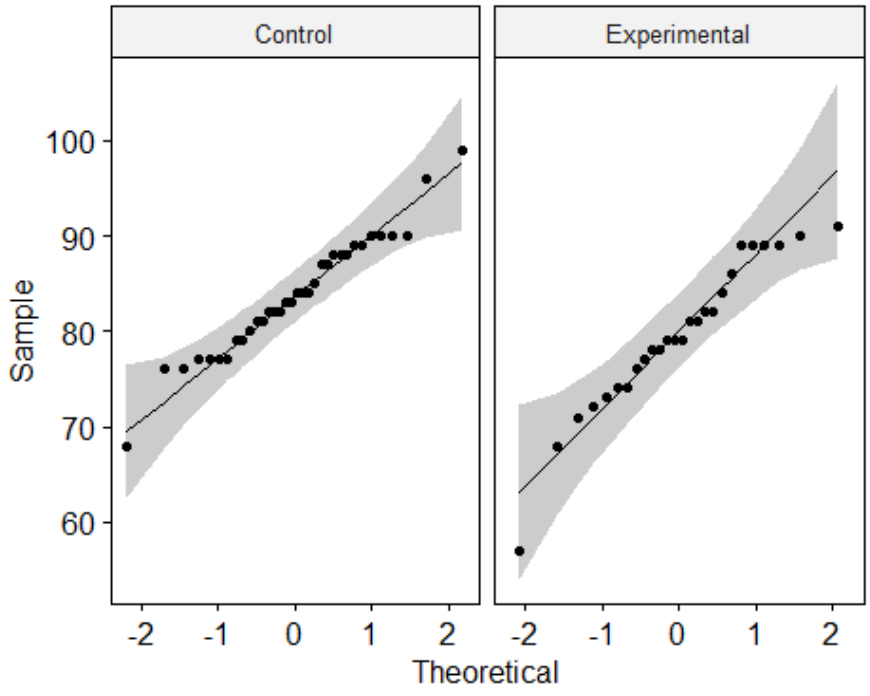
```
group        variable     n  mean    sd
  <fct>        <chr>    <dbl> <dbl> <dbl>
1 Control      posttest    34  82.7  6.93
2 Control      pretest     34  83.8  6.29
3 Experimental posttest    26  78.7  9.23
4 Experimental pretest     26  79.5  7.94
```

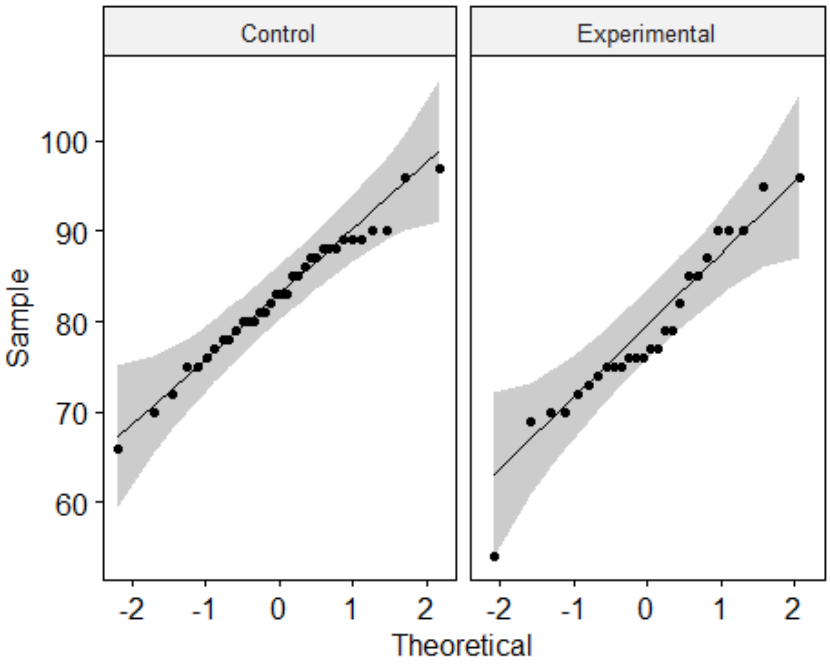



```
.y.      group1
group2          df statistic     p p.adj p.adj.signif
* <chr>    <chr>   <chr>        <dbl>     <dbl> <dbl> <dbl> <chr>
1 posttest Control Experimental    57     0.312 0.756 0.756 ns
> get_emmeans(pwc)
# A tibble: 2 x 8
  pretest group        emmean    se    df conf.low conf.high method
    <dbl> <fct>         <dbl> <dbl> <dbl>    <dbl>     <dbl> <chr>
1    81.9 Control        81.2 0.953    57     79.3      83.1 Emmeans te
st
2    81.9 Experimental   80.7 1.10     57     78.5      82.9 Emmeans te
st
```

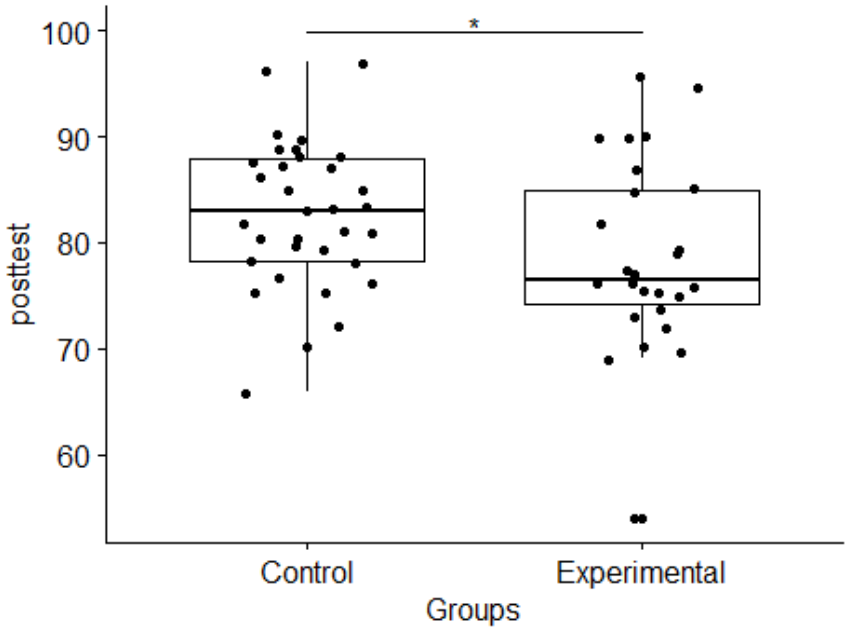

100
90
80
70
60
pretest
Control
Experimental
Groups

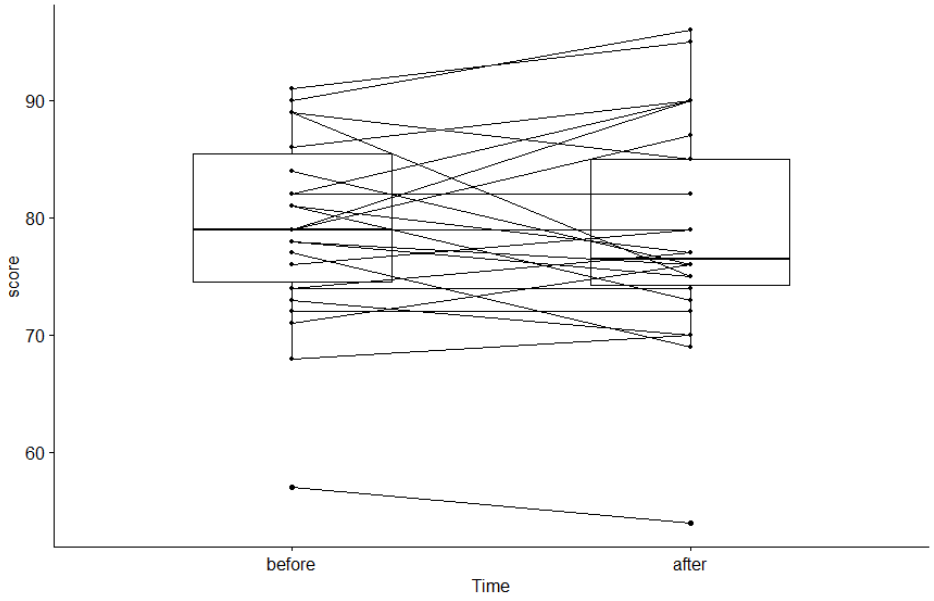

90
80
70
60
posttest
Control
Experimental
Groups

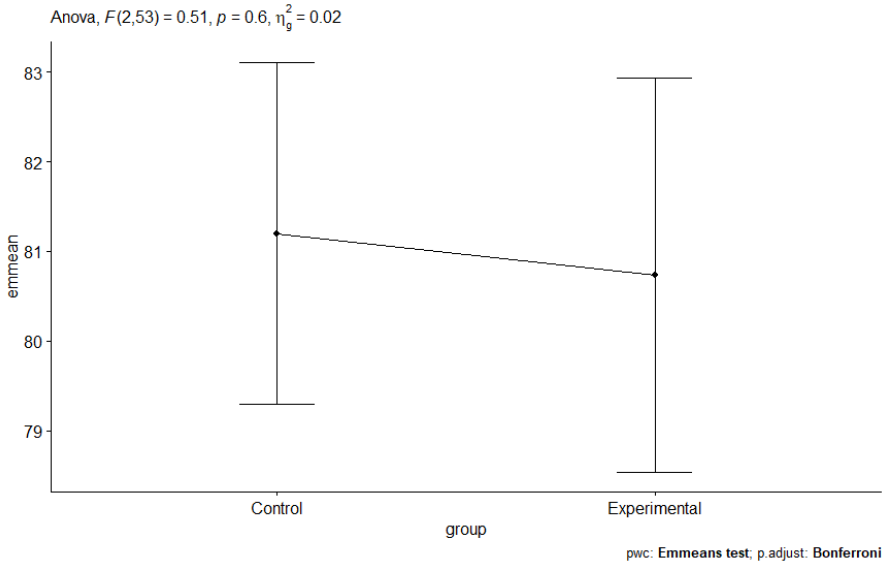

Control
Experimental
100
90
80
70
60
Sample
-2
-1
0
1
2
Theoretical

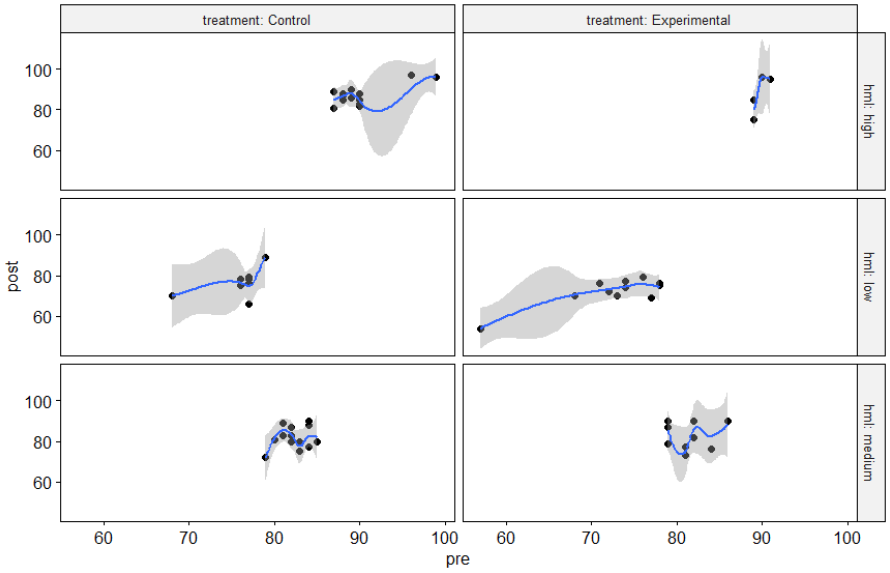


T test, $t(46.62) = 2.23$, $p = 0.031$, $n = 60$

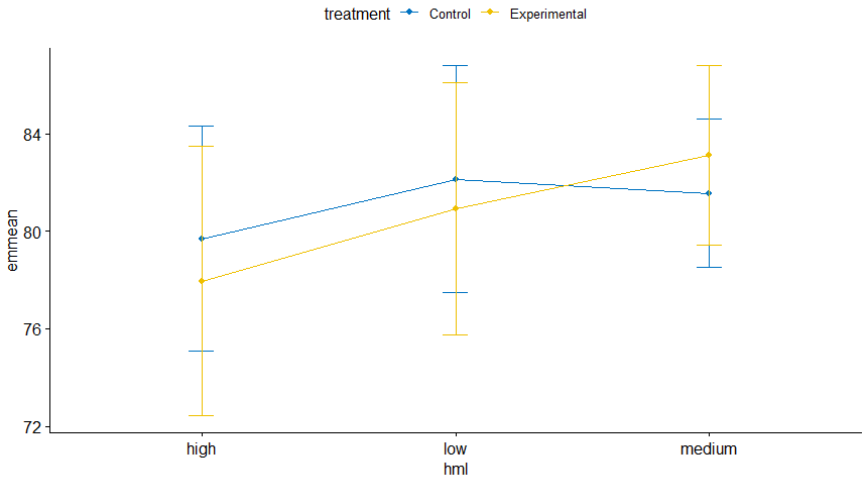



90
80
70
60
score
before
after
Time

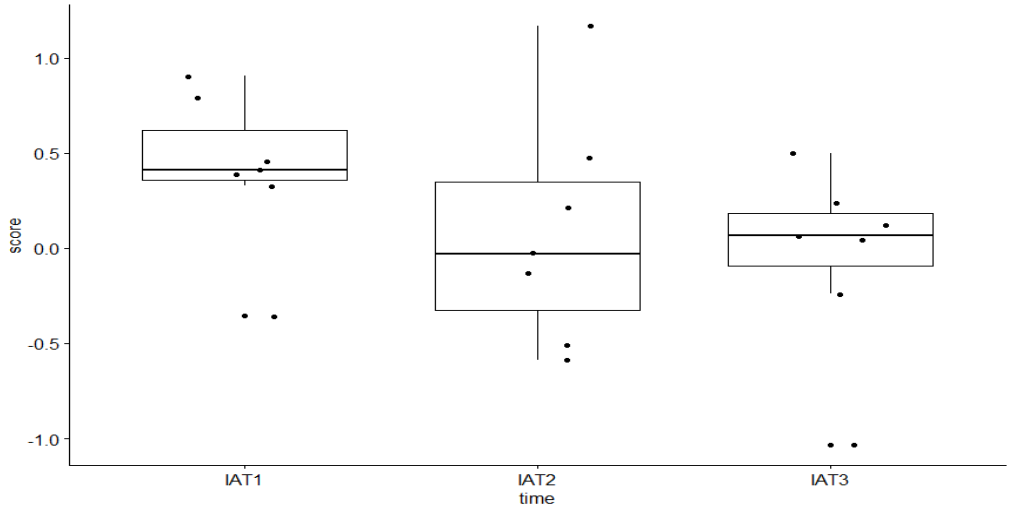

Anova, F(2,53) = 0.51, p = 0.6, η²g = 0.02
83
82
81
80
79
emmean
Control
Experimental
group
pwc: Emmeans test; p.adjust: Bonferroni

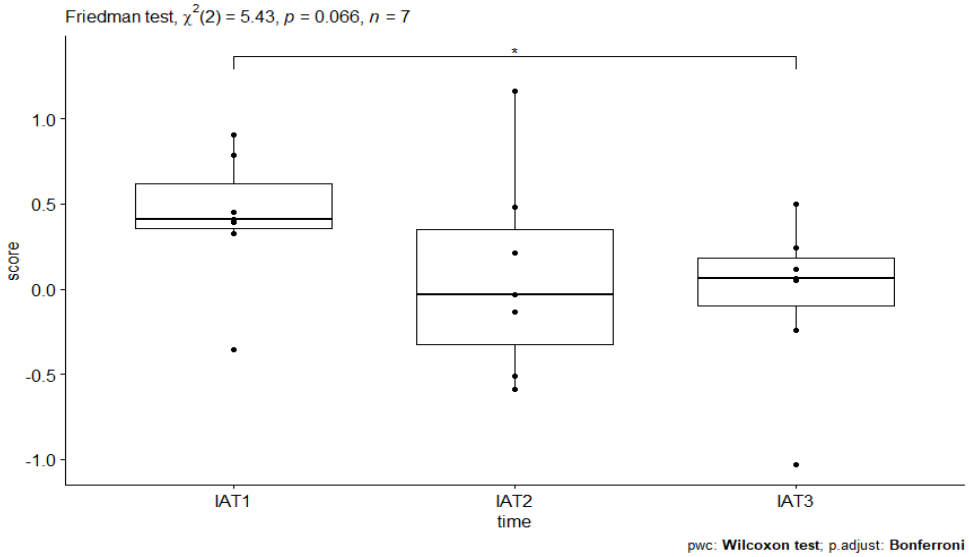


Anova, $F(2,53) = 0.51$, $p = 0.6$, $\eta_g^2 = 0.02$



pwc: **Emmeans test**; p.adjust: **Bonferroni**




Friedman test, $\chi^2(2) = 5.43$, $p = 0.066$, $n = 7$
*
1.0
0.5
0.0
-0.5
-1.0
score
IAT1
IAT2
IAT3
time
pwc: **Wilcoxon test**; p.adjust: **Bonferroni**